Université de Montréal

The Political Economy of Mental Health

Power and Interests Within a Complex System

par

Michael McCubbin

Programme de Sciences humaines appliquées
Faculté des études supérieures

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en sciences humaines appliquées

Février 1998

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

The Political Economy of Mental Health

Power and Interests Within a Complex System

présentée par:

Michael McCubbin

a été évaluée par un jury composé des personnes suivantes:

Gilles Bibeau
David Cohen
Deena White
David Rochefort
Huguette Béguin

Thèse acceptée le: 30. 04. 1998

Brief Abstract

This research aimed to contribute to the development of a critical theory explaining why the mental health system has been so immune to meaningful reform. The literature does not provide a systematic analysis of the theory and assumptions underlying a belief that the system does or can meet user "needs". This research critically analyzes the constructed proposition that the system *will* meet user needs because powerful actors have interests which converge with those of users, because users have sufficient power to affect system outcomes, or because policy interventions at the administrative and service delivery levels can compensate for failures under the prior two conditions. The thesis uses a systemic political economy analytical framework, within a policy sciences orientation, and draws upon literature from a wide variety of disciplines. With respect to the *interest convergence condition,* the thesis firmly rejects medical and market model narratives suggesting that through convergence of interests under those models the needs of users are or could be met. The best interests narrative is inapplicable to competent persons. With respect to incompetent persons, since it is unevaluable it cannot be rejected; but nor can it be accepted. With respect to the *power access condition*, it is obvious that users have virtually no power within the mental health system, hence it can be unambiguously rejected. The *policy intervention condition* will not be satisfied if the other conditions are not; policy makers are not immune to their own interests or to the power of others. The thesis concludes that the mental health system cannot be expected to meet user needs. The often noted failures in major reforms aimed at advancing user interests and increasing their power, personal and political, are indeed *inevitable* in a mental health system whose actual function and purpose — whatever they may be — are evidently at odds with the prominent *discourse* that the system is designed to meet the "needs" of mentally sick individuals. The thesis discusses in detail directions for further research, ensuing from concerns raised in the thesis.

Keywords—mental health, population health, policy, evaluation, needs, policy science, political economy, system, General System Theory, economics, politics, political science, public choice, psychiatry, community care, deinstitutionalization

Detailed Abstract

Within a teleological philosophy and General System Theory epistemological orientation, this thesis provides a critical, systemic political economy analysis of the rationality of the mental health system with respect to meeting the "needs" of users of that system. The thesis constructs and assesses theory regarding the operations and purpose of the mental health system.

This interdisciplinary thesis "by articles", completed for a Ph.D. in Sciences humaines appliquées (Applied Human Sciences) at the Université de Montréal, includes nine papers (three published, two submitted, four manuscripts). The accompanying text links the papers together in one research program and method and in assessment of a single research problem.

The thesis will interest persons concerned with mental health policy, policy science, critical and political economy analysis of social problems and institutions, and with the role in the social and policy sciences of values and ethics.

Problématique. While most analysts agree that the mental health system has been unable to achieve meaningful long-lasting reforms to improve outcomes for those considered mentally ill, the literature does not provide a systematic analysis of the fundamental rationality of the system — especially the theory and assumptions, explicit and implicit, that would underlie a belief that the mental health system is meeting or can meet the "needs" of users.

Objectives. This research aimed to contribute to the development of a critical theory explaining why the mental health system in Occidental countries has been so immune to meaningful reform, by creating a "road map" of the system in terms of the most fundamental categories of variables that could facilitate or impede reform. In the end, the aim has been to design a model such that, along with the application of value-based stances, the formulation of strategic planning for reforms of the mental health system will have something approaching a scientific basis.

Attention is paid to the micro (patient / substituted decision makers), macro (policy / politics over time), and meso (partial analysis of sub-systems and their interactions by looking at patients and psychiatrists) levels of a broadly defined historical mental health system.

It is important to note that the way the research problem is structured in this thesis does not require an *a priori* definition of needs, nor even acceptance that there *are* user needs.

Hypotheses. This research critically analyzes a constructed proposition which, if true, would provide one basis of support for the assumption of system rationality: that the system will meet user needs. The test for this proposition is that at least one of three conditions is satisfied: 1) *Interest convergence*: powerful actors have interests which converge with those of users. 2) *Power access*: users have sufficient power at the political, policy, administrative, and/or therapeutic levels to affect system outcomes. 3) *Policy intervention*: Policy interventions at the administrative and service delivery levels can compensate for failures under the prior two conditions.

Methodology. The thesis uses a systemic political economy analytical framework, within a policy sciences disciplinary orientation. Analytical methodologies from public choice, rational choice, utility, game, statistical, and decision theory, have been used in order to enable the analysis of collective and individual choices. This research has not generated new empirical data; rather it has drawn upon literature from a wide variety of disciplines in constructing and elaborating its arguments.

Results. With respect to the *interest convergence condition*, the thesis firmly rejects medical and market model narratives suggesting that through convergence of interests under those models the needs of users are or could be met. The best interests narrative is inapplicable to competent persons. With respect to incompetent persons, since it is unevaluable it cannot be rejected, but nor can it be accepted. Hence this thesis fails to accept the proposition that interests of powerful actors converge enough with those of users that user needs will be met. With respect to the *power access condition*, it is obvious that users have virtually no power within the mental health system, hence it can be unambiguously rejected. Insofar as either powerful actors in the system were to have interests convergent with users, or users themselves had power, policy intervention at the administrative and service delivery levels could be effective in further advancing user interests; otherwise interventions could not be expected to accomplish major long-lasting user-centered reform of the system. However, given the conclusions failing to accept the interest convergence condition and rejecting the power access condition, the thesis fails to accept the *policy intervention condition.*

Conclusion. Given the rejection of the power access condition and the failure to accept the interest convergence and policy intervention conditions, this thesis concludes that there are no reasonable grounds to accept a hypothesis that the mental health system meets user needs. The often noted failures in major reforms aimed at advancing user interests and increasing their power, personal and political, are indeed *inevitable* in a mental health system whose actual function and purpose — whatever they may be — is evidently at odds with the prominent *discourse* that the system is designed to meet the "needs" of mentally sick individuals.

Directions for Further Research. The thesis discusses in detail directions for further research, ensuing from concerns raised in the thesis regarding: society's faith in psychiatry; the legal system's failure to advance user rights; the paradox of helping professions incorporating empowerment principles; the justifiability of forced psychiatric treatment; the role of the pharmaceutical industry in shaping policy; the usefulness of the concept of user "needs"; intervention to enhance user competence; user wants versus user "satisfaction"; user representation of their own interests at the collective level; cooptation and evaluation regarding administrative structures; the actual meaning of "community care"; short-term planning perspectives; the independence of the research community from the systems of power and influence toward which they should be retaining a critical objectivity.

Summary of Thesis

Within a teleological philosophy and General System Theory epistemological orientation, this thesis provides a critical, systemic political economy analysis of the rationality of the mental health system with respect to meeting the "needs" of users of that system. The thesis constructs and assesses theory regarding the operations and purpose of the mental health system. It is interdisciplinary, drawing upon a variety of social sciences in order to facilitate the analysis of choices and dynamics involving individuals and collectivities.

This thesis, completed for a Ph.D. in Sciences humaines appliquées (Applied Human Sciences) at the Université de Montréal, incorporates within accompanying text three published articles, two submitted articles, and four manuscripts. The accompanying text (introduction, method, problématique, objectives, theory, methodology, analysis of results, conclusion) links the papers together in one research program and method and in assessment of a single research problem.

The thesis should be of particular interest to persons concerned with mental health policy, to specialists in policy science, to critical and political economy analysts of social problems and institutions, and to those interested in the role in the social and policy sciences of values and ethics.

Method

This research was conceived within a teleological philosophy, as described in the submitted manuscript "Free Will, Rationality, and Explanation: Toward a Teleological *Human* Science". The epistemological underpinnings of the research were provided by General System Theory, as described in the manuscript "General System Theory as a Postmodern Epistemology for the Social Sciences".

Problématique

The social problem which gave rise to this thesis is the inability of the mental health system to achieve meaningful long-lasting reforms that would improve the outcomes of the system with respect to those considered mentally ill. Difficulties in achieving major reform objectives are discussed in the article "Deinstitutionalization: The Illusion of Disillusion".

The existing literature has addressed aspects of the mental health policy system, but does not provide a systematic analysis of the fundamental rationality of the system. In particular, the literature has never addressed, systematically, the theory and assumptions, explicit and implicit, that would underlie a belief that the mental health system is meeting or can meet the "needs" of users.

Research Objectives

This research aimed to contribute to the development of a critical theory explaining why the mental health system has been so immune to meaningful reform, by creating a "road map" of the system in terms of the most fundamental categories of variables that could facilitate or impede reform. More generally, the thesis aimed:

- to develop a better understanding of the mental health sector as a system;
- to interpret system mutations and characteristics with respect to decisions made within social, political and economic contexts; and
- to generate an analytical framework which will facilitate policy development in the mental health and other fields.

The thesis applies to the mental health systems of Occidental countries, sacrificing immediate policy application in favour of theory development. Nevertheless, the aim of this research has been to provide a new way to conceive the major policy reform problems which are more or less common to all these countries, thereby hopefully facilitating within particular jurisdictions the development of precise research programs, policy proposals, and implementation strategies.

In the end, the aim has been to design a model such that, along with the application of value-based stances, the formulation of strategic planning for reforms of the mental health system will have something approaching a scientific basis.

Theory and Research Hypotheses

This research critically analyzes a constructed proposition which, if true, would provide one basis of support for the assumption of system rationality. The proposition assessed is that the system will meet user needs. The test for this proposition is that at least one of three conditions is satisfied; those conditions are argued to be exhaustive of narratives implicitly or explicitly characterizing the system as meeting user needs. The conditions are:

1) **Interest convergence condition** (*Paternalistic Model*): Powerful actors in the system have interests which converge with those of users.
2) **Power access condition** (*Political Model*): The system satisfies user needs as a result of user power to affect system outcomes — which power might be exercised in various ways throughout the system (at the political, policy, administrative, and/or therapeutic levels).
3) **Policy intervention condition** (*Bureaucratic Model*): Insofar as the system does not meet user needs — e.g., insofar as the paternalistic and political models are inadequate — policy interventions at the administrative and service delivery levels can compensate.

Attention is paid to the micro (patient / substituted decision makers), macro (policy / politics over time), and meso (partial analysis of sub-systems and their interactions by looking at patients and psychiatrists) levels of a broadly defined historical mental health system. It is important to note that the way the research problem is structured in this thesis does not require an *a priori* definition of needs, nor even acceptance that there *are* user needs.

Methodology

This thesis critically analyzes research questions deductively generated by the overall hypothesis that the mental system will meet the "needs" of users. Developed within a "policy sciences" orientation, the thesis is explicitly interdisciplinary, drawing upon various social sciences, particularly political science, economics, sociology, and social psychology.

The thesis uses a systemic political economy analytical framework, as described in the submitted manuscript "Mental Health Policy from a Systemic Perspective: Orientations for Strategic Reform". Analytical methodologies from public choice, rational choice, utility, game, statistical, and decision theory, have been used in order to enable the analysis of collective and individual choices. This research has not generated new empirical data; rather it has drawn upon literature from a wide variety of disciplines in constructing and elaborating its arguments.

Results

Power Access Condition. The thesis argues that there is no basis for an expectation that user needs will be met on the grounds of user access to power. Firstly, users have little power as seen by *direct measures,* such as actual observed influence in mental health policy and politics.

Secondly, *outcome measures* also strongly suggest that the system does not reflect user power, as is made evident in the manuscript "'Meeting the Needs of the Mentally Ill': A Case Study of the 'Right to Treatment' as Legal Rights Discourse in the U.S.A."

Thirdly, given the variety of user power disabilities described in the article "Extremely Unbalanced: Interest Divergence and Power Disparities Between Clients and Psychiatry", *users could not be expected to have meaningful power in the system.*

Interest Convergence Condition. The thesis asserts that if the interest convergence condition were to be met it would have to be on the basis of at least one of the following narratives: the *medical model,* the *market model,* and the *best interests model.*

Medical Model. The medical model is inadequate: the degree to which somatic factors are present in the etiology of psychological distress is not pertinent to the applicability of the medical model as a means of meeting user needs; decisions about medical interventions are not, in fact, medical decisions, even if they need to be informed by medical expertise. This is particularly true with respect to interventions at the *population* level, as argued in the manuscript "Population Health: A Call for Breadth (Mental Health) and Depth (Psychosocial Theory)".

Market Model. The market model is also inadequate, since conditions for the efficiency of a competitive market are not and probably could not be met with respect to services related to mental health, and given the vulnerability of many users.

Best Interests Model. The best interests model cannot be seen as applicable to competent users. However, as described in the article "Toward a Pure Best Interests Model of Proxy Decision Making for Incompetent Psychiatric Patients", the best interests criterion for decision making on behalf of incompetent persons is so vague as to be non-evaluable. Furthermore, it is paternalistic: decisions are made according to the standards of others regarding "what is good for the patient", rather than according to what the patient would have decided if temporarily competent — the *pure* best interests criterion. "Best interests" assessments might in fact incorporate non-patient interests.

When substituted decision making is evaluated with respect to a pure best interests criterion centered explicitly on the patient's own interests, based on autonomy rather than paternalistic values, it is extremely difficult to ensure that the process is valid, reliable, and unbiased in estimating the choices the patient would make if temporarily competent. Indeed, the small empirical literature in proxy decision making finds that such efforts are only fair to poor. These issues are discussed in detail in the manuscript "Error under a Pure Best Interests Model of Decision Making: Implications for the Justifiability of Forced Treatment".

In the mental health field, *de facto* or formal substituted decisions usually arise precisely because the patient has expressed treatment refusal. Given the problems in proxy decision making in general, combined with contrary expressed preferences and the emotional and therapeutic impacts on the patient of forcing treatment, circumstances in which forced treatment are justified must be extremely limited; indeed, an appropriate policy might be to prohibit forced psychiatric treatment. Nevertheless, forced treatment is common in the mental health system. Therefore, not only is a best interests criterion inapplicable for competent patients, it is inadequate for incompetent patients when evaluated according to pure best interests criteria. I suggest that the burden of proof that the needs of incompetent psychiatric patients are met via proxy decisions rests with those making such an assertion. There is little evidence in the literature to support such a position.

Policy Intervention Condition. We could not expect intervention at the administrative and service delivery levels to rectify the degree to which the system might otherwise not meet user needs, since insofar as those intervening do not have interests convergent with users, and are not themselves users, there is no basis for generalized confidence that they would act to meet user needs in any event.

Some actors, despite the generalizations in this thesis regarding interest divergence of categories of actors with users, might indeed be motivated to advance reforms meeting the interests of users rather than of others. Such intervention could potentially set off a dynamic resulting in long-lasting fundamental reform, and hence efforts in this regard are worthwhile, but observation and theory suggests that such efforts are much more likely to be extremely limited, localized, and temporary, given a powerful environment effectively hostile to the directions of such reforms.

Indeed, there are indications that responsibility for advancing and implementing well-intentioned government policy initiatives has often been placed in the hands of those with the *least* incentives to pursue them in accordance with the policies' stated objectives.

Conclusion

With respect to the *interest convergence condition,* the thesis firmly rejects medical and market model narratives suggesting that through convergence of interests under those models the needs of users are or could be met. The best interests narrative is inapplicable to competent persons. With respect to incompetent persons, since it is unevaluable it cannot be rejected, but nor can it be accepted. Hence this thesis fails to accept the proposition that interests of powerful actors converge enough with those of users that user needs will be met.

With respect to the *power access condition*, it is obvious that users have virtually no power within the mental health system, hence it can be unambiguously rejected.

Insofar as either powerful actors in the system were to have interests convergent with users, or users themselves had power, policy intervention at the administrative and service delivery levels could be effective in further advancing user interests; otherwise interventions could not be expected to accomplish major long-lasting user-centered reform of the system. However, given the conclusions failing to accept the interest convergence condition and rejecting the power access condition, the thesis fails to accept the *policy intervention condition.*

Given the rejection of the power access condition and the failure to accept the interest convergence and policy intervention conditions, this thesis concludes that there are no reasonable grounds to accept a hypothesis that the mental health system meets user needs. The often noted failures in major reforms aimed at advancing user interests and increasing their power, personal and political, are indeed *inevitable* in a mental health system whose actual function and purpose — whatever they may be — is evidently at odds with the prominent *discourse* that the system is designed to meet the "needs" of mentally sick individuals.

Directions for Future Research. The main contribution of this thesis lies in, I hope, the development of "better" questions, arising from a critical, systemic political economy analysis of the mental health sector as a complex system whose dynamics are propelled by the value-based constrained choices of individuals and groups.

The thesis discusses a number of research directions which appear important or interesting, given the issues raised, concerning:

- society's faith in psychiatric authority;
- the weakness of the legal system to advance user rights;
- the paradox of helping professions incorporating empowerment principles;
- the justifiability of policies permitting forced psychiatric treatment;

- the studiously ignored impact of the pharmaceutical industry in shaping the mental health system;
- the questionable usefulness of the concept of "user needs" in planning mental health services;
- the need to find ways toward enhancing user competence;
- determining user wants vs. satisfaction or preferences;
- the need for study of how users represent their own interests at the collective level;
- issues of cooptation and evaluation with respect to administrative structure;
- the actual shape of post deinstitutionalization "community care" and what it really means for individual users; and
- the problem of short-term perspectives in planning and maintaining reforms.

Finally, and, not of least importance, I argue the urgent need for the research community to self-consciously assess its critical objectivity, and address difficult ethical questions regarding its own independence from the systems of power and interests that have sustained a mental health system whose rationality, if any, does not lie in meeting the needs or interests of users.

Résumé de la thèse

Dans le cadre d'une philosophie téléologique et suivant une orientation épistémologique inspirée de la théorie générale des systèmes, cette thèse fournit une analyse critique d'économie politique systémique de la rationalité du système de santé mentale concernant la réponse aux "besoins" des usagers de ce système. La thèse construit et évalue une théorie concernant le fonctionnement et les fins du système de santé mentale. Interdisciplinaire, elle puise dans plusieurs sciences sociales afin de faciliter l'analyse des choix et des dynamiques impliquant les individus et les collectivités.

Cette thèse a été réalisée dans le cadre du programme de doctorat en Sciences humaines appliquées de l'Université de Montréal. Elle comprend, outre le texte d'accompagnement, trois articles publiés, deux articles soumis et quatre manuscrits. Le texte d'accompagnement (introduction, méthode, problématique, objectifs, cadre théorique, méthodologie, analyse des résultats et conclusion) fait le lien entre les neuf travaux et les intègre à la fois dans un programme et une méthode de recherche, et dans une démarche d'évaluation d'un problème de recherche unique.

La thèse présente un intérêt particulier pour les personnes en charge des politiques de santé mentale, pour les spécialistes des sciences des politiques publiques (*policy science),* pour les chercheurs qui analysent les problèmes sociaux et les institutions sous un angle critique et d'économie politique, et pour ceux qui s'intéressent au rôle des valeurs et de l'éthique dans les sciences sociales et dans les sciences des politiques publiques.

Méthode

Cette recherche a été conçue dans le cadre d'une philosophie téléologique, telle que décrite dans le manuscrit intitulé "Free Will, Rationality, and Explanation: Toward a Teleological *Human* Science" (soumis pour publication). Les fondements épistémologiques de la recherche proviennent de la théorie générale des systèmes, telle que décrite dans le manuscrit intitulé "General System Theory as a Postmodern Epistemology for the Social Sciences".

Problématique

Le problème social qui a généré cette thèse est celui de l'incapacité du système de santé mentale d'instaurer avec succès des réformes significatives et durables, susceptibles d'améliorer ses résultats en ce qui concerne les personnes considérées mentalement malades. Les difficultés à atteindre les objectifs des réformes sont présentées et commentées dans l'article "Deinstitutionalization: The Illusion of Disillusion".

Les écrits existants ont abordé de nombreux aspects des politiques en santé mentale, mais ne fournissent pas d'analyse systématique de la rationalité fondamentale du système. En particulier, les écrits n'ont jamais abordé de manière systématique la théorie et les présupposés, explicites et implicites, sous-jacents à la croyance selon laquelle le système de santé mentale répond ou peut répondre aux "besoins" des usagers.

Objectifs de la recherche

Cette recherche vise à contribuer au développement d'une théorie critique expliquant pourquoi le système de santé mentale a été si résistant à une réforme significative. L'explication fournie ici procède par la création d'une "carte routière" du système, laquelle indique les catégories de variables les plus fondamentales et susceptibles de faciliter ou empêcher une réforme. Plus généralement, la thèse vise à:

- développer une meilleure compréhension du secteur de la santé mentale en tant que système;
- interpréter les mutations et les caractéristiques du système en ce qui concerne les décisions prises à l'intérieur de contextes sociaux, politiques et économiques; et
- générer un cadre analytique qui facilitera le développement de politiques dans le champ de la santé mentale et dans d'autres champs.

Cette thèse s'applique aux systèmes de santé mentale des pays occidentaux, sacrifiant l'application immédiate de politiques en faveur du développement de la théorie. Il reste que le but était de fournir une manière nouvelle de concevoir les principaux problèmes en matière de réforme des politiques — problèmes plus ou moins communs à ces pays — et de faciliter ainsi le développement de programmes précis de recherche, de propositions de politiques et de stratégies d'implantation, le tout à l'intérieur de juridictions particulières.

Enfin, l'objectif était de concevoir un modèle où, conjointement à l'application de prises de position fondées sur des valeurs, la formulation de la planification stratégique pour les réformes du système de santé mentale disposera de fondements se rapprochant d'une base scientifique.

Hypothèses théoriques et de recherche

Cette recherche analyse de manière critique une proposition construite qui, si elle est validée, fournirait une base soutenant le présupposé de la rationalité du système de santé mentale. La proposition ainsi évaluée est à l'effet que le système répondra aux besoins des usagers. Le test de cette proposition est fondé sur trois conditions, dont au moins une doit être satisfaite; ces conditions sont présentées ici comme exhaustives, dans le sens où elles rendent compte des différents discours (*narratives*) où le système est implicitement ou explicitement caractérisé comme répondant aux besoins des usagers. Ces conditions sont:

1) **La convergence des intérêts** (*Modèle paternaliste)*: Les acteurs ayant du pouvoir dans le système ont des intérêts qui convergent avec ceux des usagers.
2) **L'accès au pouvoir** (*Modèle politique*): Le système satisfait les besoins des usagers en raison du pouvoir qu'ont ceux-ci d'en affecter les résultats — un pouvoir qui peut être exercé de différentes manières à travers le système (aux niveaux politique, administratif, thérapeutique ou des politiques publiques).
3) **L'intervention par les politiques publiques** (*Modèle bureaucratique*): Dans la mesure où le système ne répond pas aux besoins des usagers — par exemple, dans la mesure où les modèles paternaliste et politique sont inadéquats — les interventions par des politiques touchant le niveau administratif et la dispensation de services peuvent compenser.

L'analyse tient compte des niveaux micro (patient / curateurs), macro (politiques publiques / dimension politique à travers le temps) et meso (analyse partielle des sous-systèmes et de leurs interactions par un regard porté sur les patients et les psychiatres) d'un système de santé mentale qui existe et évolue dans le temps. Il importe de noter que la manière dont le problème de recherche est structuré dans cette thèse ne requiert par une définition *a priori* des besoins des usagers, non plus que l'acceptation de l'idée que de tels besoins *existent* effectivement.

Méthodologie

Cette thèse analyse de manière critique les questions de recherche générées de manière déductive par l'hypothèse générale selon laquelle le système de santé mentale répondra aux "besoins" des usagers. Développée suivant l'orientation des *policy sciences*, la thèse est explicitement interdisciplinaire, puisant à une variété de sciences sociales, en particulier la science politique, la science économique, la sociologie et la psychologie sociale.

Elle utilise un cadre d'analyse d'économie politique systémique, tel que décrit dans le manuscrit "Mental Health Policy from a Systemic Perspective: Orientations for Strategic Reform" (soumis pour publication). Les méthodologies analytiques provenant des théories du choix public, du choix rationnel, de l'utilité, du jeu, des statistiques et de la décision ont été utilisées afin de permettre une analyse des choix collectifs et individuels. Cette recherche n'a pas généré de données empiriques inédites; elle a plutôt puisé dans les écrits provenant d'une grande variété de disciplines afin de construire et élaborer ses arguments.

Résultats

La condition de la convergence des intérêts. La thèse affirme que si la condition de la convergence des intérêts était satisfaite, ce serait sur la base d'au moins un des discours suivants: le *modèle médical*, le *modèle du marché* et le *modèle des meilleurs intérêts*.

Le modèle médical. Le modèle médical est inadéquat: l'importance actuelle ou potentielle de facteurs somatiques dans l'étiologie de la détresse psychologique n'est pas pertinente à l'application du modèle médical comme moyen de répondre aux besoins des usagers; les décisions concernant les interventions médicales ne constituent pas, en fait, des décisions strictement médicales, même si elles ont besoin d'être informées par l'expertise médicale pour être mises en oeuvre. Ceci est particulièrement vrai en ce qui concerne les interventions au niveau des populations, comme le montre la discussion présentée dans le manuscrit "Population Health: A Call for Breadth (Mental Health) and Depth (Psychosocial Theory)".

Le modèle du marché. Le modèle du marché est inadéquat, parce que les conditions mêmes de l'efficience d'un marché compétitif ne sont pas — ni en théorie ni dans les faits — satisfaites en ce qui concerne les services de santé mentale; la vulnérabilité de plusieurs usagers empêche également l'applicabilité de ce modèle.

Le modèle des meilleurs intérêts. Le modèle des meilleurs intérêts est inapplicable dans le cas d'usagers compétents. L'article "Toward a Pure Best Interests Model of Proxy Decision Making for Incompetent Psychiatric Patients", montre toutefois que même lorsqu'il est utilisé pour la prise de décision au nom des personnes jugées incompétentes, les critères dérivés de ce modèle sont si vagues que le modèle n'est pas évaluable. De surcroît, ces critères sont paternalistes: les décisions sont prises en fonction des standards d'autrui à propos de "ce qui est bon pour le patient", plutôt qu'en fonction de ce que le patient aurait décidé s'il avait été temporairement compétent — soit le critère du meilleur intérêt *pur*. Dans les faits, les évaluations fondées sur les "meilleurs intérêts" sont susceptibles d'incorporer des intérêts autres que ceux des patients.

Quand la prise de décision substituée est évaluée selon le critère du meilleur intérêt pur, centré explicitement sur les intérêts propres au patient et fondé sur des valeurs d'autonomie plutôt que des valeurs paternalistes, il est extrêmement difficile d'assurer que le processus est valide, fiable et non-biaisé dans l'estimation des choix que le patient aurait fait s'il avait été temporairement compétent. En effet, la (peu abondante) littérature empirique en matière de prise de décision par procuration a montré que de tels efforts donnent des résultats assez décevants. Ces questions sont discutées en détail dans le manuscrit "Error under a Pure Best Interests Model of Decision Making: Implications for the Justifiability of Forced Treatment".

Dans le champ de la santé mentale, les décisions substituées *de facto* ou formelles se produisent d'habitude précisément parce que le patient a exprimé un refus de traitement. Étant donné les problèmes dans la prise de décision par procuration en général, combinés à l'expression de préférences contraires et aux impacts émotionnels et thérapeutiques du traitement forcé sur le patient, les circonstances dans lesquelles le traitement forcé est justifié doivent être extrêmement restreintes; en fait, la prohibition pure et simple du traitement forcé pourrait s'avérer une politique appropriée. Néanmoins, le traitement forcé est pratique courante dans le système de santé mentale. Ainsi, non seulement le critère des meilleurs intérêts est-il inapplicable pour les patients compétents, mais il est aussi inadéquat pour les patients incompétents quand il est évalué selon les critères du meilleur intérêt pur. J'avance que le fardeau de la preuve à l'effet que les besoins des patients psychiatriques incompétents sont rencontrés via les décisions par procuration repose sur ceux-là mêmes qui font cette affirmation. La littérature pertinente donne peu de preuves à l'appui d'une telle position.

Condition de l'accès au pouvoir. La thèse soutient qu'il n'y a aucune assise permettant de supposer que les besoins des usagers seront comblés sur la base de leur accès au pouvoir. Premièrement, les usagers ont peu de pouvoir observable par des *mesures directes*, telles que l'influence concrète et visible au niveau politique et dans les politiques de santé mentale.

Deuxièmement; les mesures des résultats du système nous suggèrent elles aussi fortement que celui-ci ne reflète pas le pouvoir des usagers, comme il est démontré de manière évidente dans le manuscrit "'Meeting the Needs of the Mentally Ill': A Case Study of the 'Right to Treatment' as Legal Rights Discourse in the U.S.A."

Troisièmement, étant donné la variété des désavantages et handicaps affectant le pouvoir des usagers — exposée dans l'article intitulé "Extremely Unbalanced: Interest Divergence and Power Disparities Between Clients and Psychiatry" — *on ne peut s'attendre* à ce que les usagers aient un pouvoir significatif dans le système.

La condition de l'intervention par les politiques. Nous ne pouvons pas nous attendre à ce que l'intervention aux niveaux de l'administration et de la dispensation des services rectifie les circonstances qui font que le système ne répond pas aux besoins des usagers: dans la mesure où ceux qui interviennent n'ont pas d'intérêts convergents avec ceux des usagers et ne sont pas eux-mêmes usagers, rien ne permet de croire qu'ils agiraient de manière à répondre aux besoins de ces derniers.

Malgré les généralisations posées dans cette thèse quant à la divergence entre les intérêts de certaines catégories d'acteurs et ceux des usagers, il est possible que certains acteurs soient motivés à mettre de l'avant des réformes rencontrant les intérêts des usagers plutôt que ceux d'autres groupes. De telles interventions pourraient potentiellement déclencher une dynamique conduisant à une réforme fondamentale et durable, et dès lors les efforts dans ce sens sont valables. Cependant, l'observation empirique et la théorie suggèrent que de tels efforts sont bien davantage susceptibles d'être limités, localisés et temporaires, étant donné l'existence d'un environnement puissant qui est effectivement hostile aux orientations de telles réformes.

En effet, certains faits indiquent que la responsabilité de l'avancement et de l'implantation de politiques gouvernementales bien-intentionnées a été souvent confiée à ceux-là mêmes qui ont le *moins* d'incitatifs à les mettre en oeuvre dans le respect de leurs objectifs déclarés.

Conclusion

En ce qui concerne la *condition de la convergence des intérêts*, la thèse rejette fermement le discours du modèle médical et celui du modèle du marché, lesquels suggèrent qu'à travers la convergence des intérêts, les besoins des usagers sont ou pourraient être comblés. Quant au discours des meilleurs intérêts, il ne s'applique pas aux personnes compétentes. Dans le cas des personnes incompétentes, ce discours ne peut être rejeté, mais il ne peut non plus être accepté car il n'est pas évaluable. Ainsi, la thèse ne peut accepter la proposition à l'effet que les intérêts des acteurs disposant de pouvoir convergent suffisamment avec ceux des usagers pour que les besoins de ces derniers soient rencontrés.

En ce qui concerne la *condition de l'accès au pouvoir*, il est évident que les usagers n'ont virtuellement aucun pouvoir dans le système de santé mentale; par conséquent, cette condition peut être rejetée sans équivoque.

Dans la mesure où les acteurs disposant de pouvoir dans le système auraient des intérêts convergents avec ceux des usagers, ou encore que les usagers aient eux-mêmes du pouvoir, l'intervention par des politiques aux niveaux de l'administration et de la dispensation des services pourrait être efficace à faire avancer les intérêts des usagers. Autrement, on ne pourrait s'attendre à ce que les interventions accomplissent des réformes majeures et durables, centrées sur les usagers, dans le système. Or, étant donné que les conclusions ne permettent pas d'accepter la condition de la convergence des intérêts et que, d'autre part, elles conduisent au rejet de la condition de l'accès au pouvoir, la thèse ne peut non plus retenir la *condition de l'intervention par des politiques*.

Étant donné le rejet de la condition de l'accès au pouvoir et l'impossibilité d'accepter les conditions de la convergence des intérêts et de l'intervention par des politiques, cette thèse conclut qu'il n'existe aucun fondement raisonnable permettant d'accepter une hypothèse voulant que le système de santé mentale réponde aux besoins des usagers. Les échecs fréquemment soulignés des réformes majeures visant à promouvoir les intérêts des usagers et augmenter leur pouvoir, tant au niveau personnel que politique, sont en effet inévitables dans un système de santé mentale dont les fonctions et les fins véritables — quelles qu'elles soient — contredisent de tout évidence un discours prédominant, à savoir que le système est conçu de manière à combler les "besoins" des individus mentalement malades.

Orientations pour la recherche future. Une principale contribution de cette thèse devrait résider dans le développement de "meilleures" questions, résultant d'une analyse critique, combinant la perspective systémique et l'économie politique, du secteur de la santé mentale abordé en tant que système complexe dont les dynamiques sont catalysées par les choix — à la fois contraints et fondés sur des valeurs — des individus et des groupes.

Cette thèse expose un certain nombre d'axes de recherche qui paraissent importants ou intéressants en regard des questions soulevées, concernant:

- la foi de la société dans l'autorité psychiatrique;
- la faiblesse du système légal à faire avancer les droits des usagers;
- le paradoxe des professions d'aide cherchant à incorporer des principes d'*empowerment* de leurs clientèles;
- la justifiabilité des lois et règlements permettant le traitement psychiatrique forcé;
- l'impact soigneusement ignoré de l'industrie pharmaceutique sur le système de santé mentale;

- l'utilité douteuse du concept de "besoins" des usagers dans la planification des services de santé mentale;
- le besoin de trouver des manières de faire menant à l'amélioration des compétences des usagers;
- la détermination des désirs réels des usagers versus leur satisfaction ou leurs préférences;
- comment les usagers représentent et revendiquent leurs intérêts au niveau collectif;
- des questions touchant la cooptation et l'évaluation en ce qui concerne la structure administrative;
- la forme concrète des "services communautaires" et ce que ceci signifie vraiment pour les usagers individuels; et
- le problème posé par des perspectives à court terme dans la planification et le maintien des réformes.

Finalement, et ceci n'est pas de la dernière importance, cette thèse soutient qu'il existe un besoin urgent, pour la communauté des chercheurs, d'évaluer consciemment son objectivité critique et d'aborder des questions éthiques difficiles concernant son propre indépendance par rapport aux systèmes de pouvoir et d'intérêts qui ont par ailleurs soutenu un système de santé mentale dont la rationalité, si tant est qu'elle existe, ne se trouve pas dans la poursuite des besoins ou des intérêts des usagers.

Table of Contents

List of Figures

Acknowledgments

I would like to thank all of those persons who contributed in various ways to the development of the papers in this thesis; they are named in the acknowledgments for each particular paper.

I wish to express here my warm appreciation to those who have contributed in important ways to my overall success in the doctoral program and in completing this thesis.

It is expected that every thesis acknowledge the thesis advisor. However, my appreciation for the help I have received from David Cohen goes far beyond the routine: because of Dr. Cohen I improved my writing enormously. He would never accept fuzzy writing, nor the fuzzy thinking that such writing often tries to mask. He imparted to me some of the art and science of writing articles for peer reviewed journals — a very difficult undertaking indeed. I have no doubt that without his advice I would never have achieved publication for my work.

I also appreciate that Dr. Cohen encouraged me to follow my own thinking in my own directions — even when I wasn't sure where I was going. His confidence in me provided a solid anchor during a long period of doctoral studies when many students lose their way because of lack of self-confidence and fear of going "too far". Dr. Cohen never worried about me going too far; he just tried to ensure that I could figure out approximately where I was when I got there.

Dr. Cohen has helped me immensely in finding the financial resources necessary to complete the doctoral program, and always had time to advise on the many little details about the nature of the academic "system" and how to construct a career in it.

I simply cannot overemphasize how invaluable Dr. Cohen was for me in accomplishing this thesis and arriving at the end of a process which while arduous was extremely fulfilling and stimulating.

During the whole period of my doctoral studies I have been fully integrated at GRASP, the Groupe de recherche sur les aspects sociaux de la santé et de la prévention. I arrived at GRASP from Ontario, speaking very poor French, and with an educational background quite different from others at GRASP. Nevertheless, GRASP made me feel at home and provided one of the best environments for completing doctoral studies a student could have. The norms of GRASP, in terms of scientific production and quality, are very high — students often do not appreciate how beneficial such an environment is, and particularly that of GRASP ... until they visit another.

I want to thank all of the GRASP students, staff, research professionals and researchers for their help, encouragement, and friendliness that have made "coming to work" every day a pleasure. I especially want to thank Jacqueline Bacon, Jocelyn Bisson, Jocelyne Boivin, Bernadette Dallaire, and Line Garand, for their help and advice over the years, and especially for their important role in creating a human environment.

I have undoubtedly consumed far more resources at GRASP than the usual doctoral student who has worked much more on his thesis than on contract work. Besides the use of a telephone, office and computer, I have benefited from extensive use of the photocopier and mail service — of crucial importance for a thesis based on a large amount of literature. I hope in the end I will have compensated GRASP in some way for the enormous benefits I have received by doing my doctoral research there.

I have received motivating encouragement and helpful advice from a number of researchers in addition to David Cohen. Among them I particularly want to thank: Andrée Demers, Marc Renaud, and Henri Dorvil, all at GRASP; Frédéric Lesemann, former head of my doctoral program in Sciences humaines appliquées, and now Directeur, Institut National de la Recherche Scientifique — Culture et Société; and David Weisstub, Philippe Pinel Professor of Legal Psychiatry and Biomedical Ethics, Faculté de médecine. Because these people conveyed to me *their* confidence in what I could do, I found myself continually crossing the boundaries of what *I* had thought I could do.

I thank the Université de Montréal for the generous bursaries they provided to me over several years; they made the difference between starvation and getting by, and motivated me to maintain the standards of quality I had shown in my work, and surpass them.

A doctoral thesis in a sense encapsulates a long period of education and life experience extending far in the past — a thesis is the more visible tip of the iceberg that is a well-trained scientist. I want to thank persons who in the past had helped make possible the eventual accomplishment of a Ph.D., notably: Louis Rouillard, my supervisor during most of the years I worked at Health and Welfare Canada — he helped developed my professionalism and was supportive in my plans to return to university; Elizabeth Carlson, Dept. of Geography, Ryerson Polytechnic University; Ian Greene, Public Policy and Administration Program, York University; and Harvey Simmons, Dept. of Political Science, York University.

Finally, but not of least importance, I want to thank a special person — she knows who she is — who over the last three years has been a valuable colleague as well as help-mate. Her support in many ways has been indispensable to bringing this thesis to fruition. I also want to acknowledge the sacrifices made by my two nieces, Anysia and Katrina; they spent their earliest years almost never seeing their uncle because of the work and especially the financial demands of a Ph.D. program. They are still too young to understand why; I do hope that someday my example will help encourage them to strive toward the fulfillment that education can bring.

Dedication

I dedicate this thesis to my father, Lcdr. Ernest McCubbin. His values of hard work and uncompromising personal integrity have been a crucial inspiration to me and shaped how my life has unfolded. Coming from a childhood of extreme poverty, even by the standards of the Great Depression which devastated the Canadian prairies where he grew up, he believed in the importance of education, eventually achieving a master's degree at M.I.T. in aeronautics and astronautics. He always inquired as to what was left unsaid, behind the superficial wisdom of common beliefs. He was a true intellectual as well as a scientist, yet very modest. He never allowed success nor systems of power and influence to invade and distort his own view of the world and how it should be. He was, like many of his generation, a man of few, but very well chosen words. I know that if he were here today he would say a few words to express his pride in what I have accomplished.

The Political Economy of Mental Health

Power and Interests Within a Complex System

Introduction

This thesis addresses an important social problem: the inability of the mental health system to achieve long-lasting reforms that would improve outcomes with respect to those considered mentally ill. Conceived within a teleological philosophy whereby understanding human phenomena requires focusing on the purposive choices of individuals and collectivities, this research uses a General System Theory (GST) epistemological and analytical framework to progress toward a critical theory which would explain why the system has been so immune to meaningful reform. The fundamental variables of the system are incorporated within a broad enough scope to reveal long term patterns and strategies. Viewing the mental health system as developmentally dynamic provides a robust model to generate hypotheses about the system's nature, actor motivations, and system change.

While there has been no shortage of critiques of system processes or manifestations such as asylums and psychiatric practice, the literature has not methodically analyzed the set of conditions necessary or sufficient to support the expectation or assumption of system rationality with respect to user needs. This research critically analyzes a constructed proposition which, if true, would provide one basis of support for the assumption of system rationality. The proposition assessed is that the system will meet user needs. The test for this proposition is that at least one of three conditions is satisfied; those conditions are argued to be exhaustive of narratives implicitly or explicitly characterizing the system as meeting user needs.

The conditions are:

1) *Interest convergence condition*: Powerful actors in the system have interests which converge with those of users.
2) *Power access condition*: The system satisfies user needs as a result of user power to affect system outcomes — which power might be exercised in various ways throughout the system (at the political, policy, administrative, and/or therapeutic levels).
3) *Policy intervention condition*: Insofar as the system does not meet user needs — e.g., insofar as the paternalistic and political models are inadequate — policy interventions at the administrative and service delivery levels can compensate.

Attention is paid to the micro (patient / substituted decision makers), macro (policy / politics over time), and meso (partial analysis of sub-systems and their interactions by looking at patients and psychiatrists) levels of a broadly defined historical mental health system. It is important to note that the way the research problem is structured in this thesis does not require an *a priori* definition of needs, nor even acceptance that there *are* user needs.

This thesis critically analyzes research questions deductively generated by the above conditions. Developed within a "policy sciences" orientation, the thesis is explicitly interdisciplinary, drawing upon various social sciences, particularly political science, economics, sociology, and social psychology. The thesis should be of particular interest to persons concerned with mental health policy, to specialists in policy science, to critical and political economy analysts of social problems and institutions, and to those interested in the role in the social and policy sciences of values and ethics.

The thesis uses a systemic political economy analytical framework. Analytical methodologies from public choice, rational choice, utility, game, statistical, and decision theory, have been used in order to enable the analysis of collective and individual choices. This research has not generated new empirical data; rather it has drawn upon literature from a wide variety of disciplines in constructing and elaborating its arguments. The analysis, while primarily taking the form of theoretical or logical argument, will draw upon publications from various disciplines and from public sources.

The thesis applies to the mental health systems of Occidental countries, sacrificing immediate policy application in favour of theory development.

Nevertheless, the aim of this research has been to provide a new way to conceive the major policy reform problems which are more or less common to all these countries, thereby hopefully facilitating within particular jurisdictions the development of precise research programs, policy proposals, and implementation strategies.

This is a "thesis by articles", incorporating 9 separate papers. Three have been published, two submitted for publication, and four are manuscripts subject to further revision prior to submission to a journal. These papers were developed within the context of various stages of developing the thesis. The various articles each incorporate the usual elements of a scientific article which is in the nature of a theoretical review: problématique, method/methodology, propositions/theory, arguments incorporating literature review. However, within the development of the thesis they each have tended to play a more or less specialized role. Hence they will be introduced and summarized in the sections of the thesis to which they are most pertinent, but will also be referred to elsewhere.

However, in order to facilitate reference to the papers, several of which are referred to in various places in the thesis, they have been attached at the end of the thesis. Note that the reference list in the text accompanying the articles includes only entries referred to in that text; each paper includes its own reference list. For ease of reference the papers are ordinarily referred to by short versions of the titles.

The following papers are incorporated in this thesis, listed in the order that they will be formally introduced in the thesis, preceded by the short title and followed by the chapter in which they are formally introduced ("MM" refers to myself as author):

The Free Will paper: MM (1997c). Free will, rationality, and teleological explanation: Toward a teleological *human* science. Submitted to *Journal of Mind and Behavior*, 29 Nov. 1997. (Method — Ontology)

The GST paper: MM (1997d). General System Theory as a postmodern epistemology for the social sciences. Unsubmitted manuscript. (Method — Epistemology)

The Deinstitutionalization paper: MM (1994b). Deinstitutionalization: The illusion of disillusion. *Journal of Mind and Behavior, 15,* 35-53. (Problématique)

The Strategic paper: MM, and Cohen, D. (1997a). Mental health policy from a systemic perspective: Orientations for strategic reform. Submitted to *Journal of Health Politics, Policy and Law,* 25 Nov. 1997. (Methodology)

The Unbalanced paper: MM, and Cohen, D. (1996). Extremely unbalanced: Interest divergence and power disparities between clients and psychiatry. *International Journal of Law and Psychiatry, 19,* 1-25. (Results)

The Pop-Health paper: MM (1997b). Population health: A call for breadth (mental health) and depth (psychosocial theory). Unpublished manuscript. (Results — Interest Convergence — Medical Model)

The Pure Best Interests paper: MM, and Weisstub, D.N. (1998a). Toward a pure best interests model of proxy decision making for incompetent psychiatric patients. *International Journal of Law and Psychiatry, 21,* 1-30. (Results — Interest Convergence — Best Interests Model)

The Error paper: MM (1997a). Error under a pure best interests model of proxy decision making: Implications for the justifiability of forced treatment. Unsubmitted manuscript, presented to the XXIInd International Congress on Law and Mental Health, Montréal, June 1997. (Results — Interest Convergence — Best Interests Model)

The Right to Treatment paper: MM, and Weisstub, D.N. (1998b) "Meeting the needs of the mentally ill": A case study of the "right to treatment" as legal rights discourse in the U.S.A. (Results — Power Access)

This thesis distinguishes between "method" and "methodology"; the former, including my basic ontology and epistemological orientations, is discussed prior to the problématique, research objectives, theory and hypotheses, since the selection of research problem and how to view it was a consequence of my method.

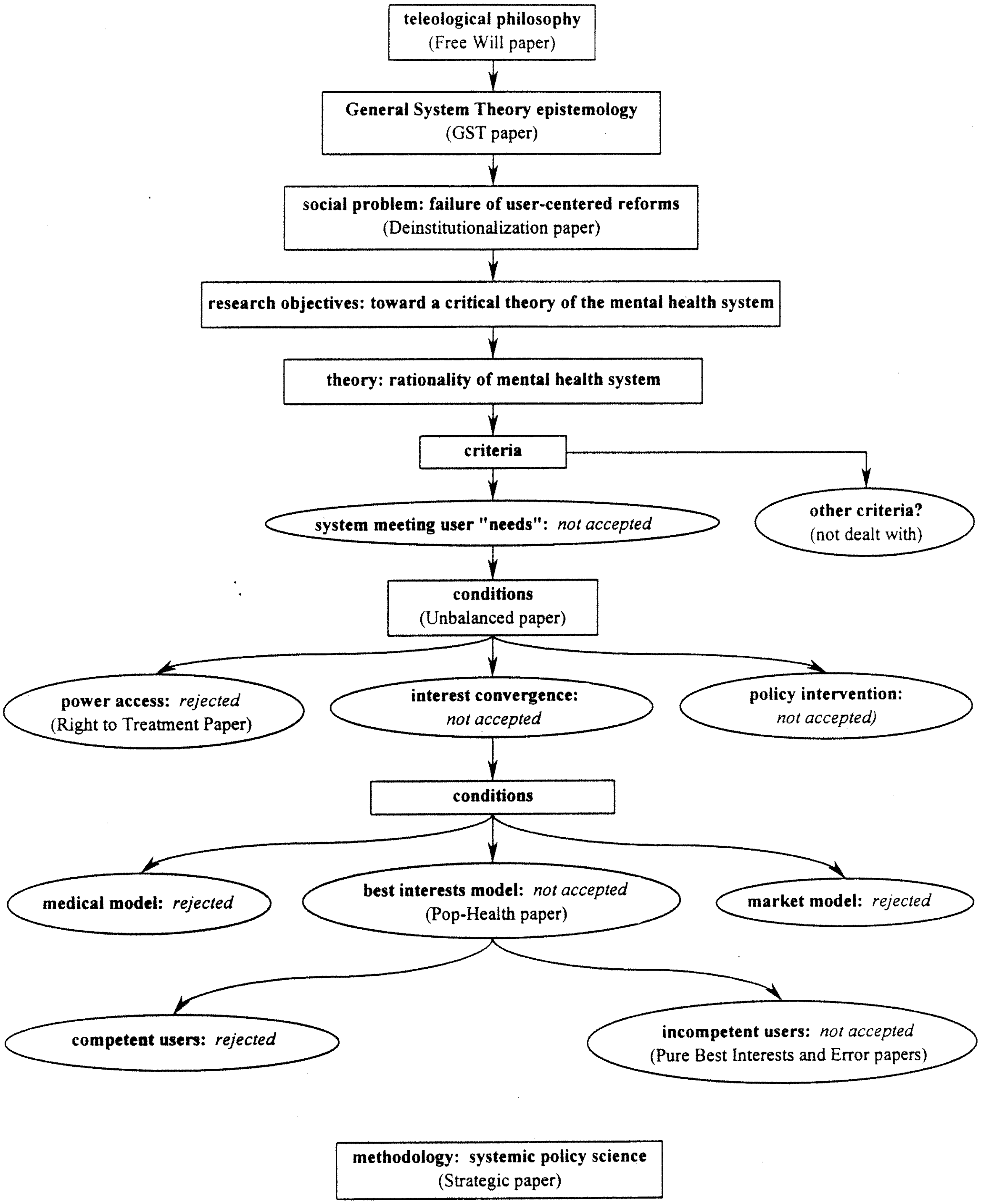
This is a very long thesis, in that it includes nine papers along with accompanying text. Those papers, designed to address more narrowly defined problems than the overall research problem of this thesis, and designed to be of interest to different audiences, will contain some discussions and details which could distract the reader trying to follow the logic of the overall thesis. Hence a useful approach to reading this thesis might be to first read through the accompanying text quickly, without referring to the articles (except that the Unbalanced paper should be read when it is introduced in the Results section); the text summarizes and advances what in the articles is most pertinent for the overall arguments.

Indeed, in order that the text can almost stand alone in coherently advancing this research project, some of it repeats sections of some articles (particularly the problématique). In a subsequent more careful reading, the papers could then be read or skimmed through. A final quick reading of the text alone will help to clarify in the reader's mind the overall structure and directions of this thesis.

Finally, the reader would be advised to refer to the thesis summary, and to the following outline of the structure of the thesis arguments which appears on the next page.

Schema: Thesis Theory And Its Context

(dispositions of propositions indicated in italics)



Method: Teleological Ontology and GST Epistemology

While unfortunately the "method" behind most positivist research, perhaps more so quantitative than qualitative, is largely left implied by the methodological processes enumerated, a theoretical thesis especially requires explicit explanation as to what in general qualifies as "data" — emphasizing *concepts* rather than "facts" — and how, independent of methodologies contingent upon particular research problems, such data will be approached, developed and organized.

To distinguish this from methodology, I will define "method" as it is understood in this thesis[1] as designating the fundamental ways in which a researcher perceives reality or the objects of observation as interrelated (or not) with surrounding phenomena. Method understood this way would be (or should be) derived from the researcher's world-view, life experience, and general philosophical attitudes and values. While a researcher's method may be in a dialectic with new experiences and activities, including specific research projects and programs, it can for the most part be viewed as providing an important part of the context within which is situated the selection of particular research objects and the methodologies used to observe them.

Ontology: Free Will and its Implications for *Human* Social Sciences

As I understand it here, method provides a bridge between the basic values of the researcher and the norms of science. Where a method touches basic values most closely is at the ontological underpinnings of that method: metaphysical assumptions and beliefs which define the boundaries of what the researcher could

---

[1] The terms "method" or "méthode" have varying senses among different academic cultures and approaches. Most such senses place method somewhere between ontology and methodology, often in the sense of a meta-methodology or analytical approach not specific to a particular research object.

accept or treat as "real" — and, indeed, defines for the researcher the meaning for him or her of that concept[2].

Similarly, the degree to which one can *define* one's ontology is likely quite limited, due to the philosophical problem known as the "incompleteness theorem": full explanation of any phenomenon requires abstraction beyond that phenomenon (von Bertalanffy, 1981). Yet once we reach ontological assumptions, there is little upon which we can base a *rational* explanation — providing the "why" in reasoned terms[3] — beyond faith, pure values, preferences and, given those, perhaps pragmatism as well. One cannot appeal to reality in order to define what is real.

---

2. Unless indicated otherwise, references in this thesis to "reality" mean only that which we commonly view or assume as "real", and does not imply a particular ontological stance about the nature or existence of reality. My own view of "reality" is that whether anything is "real" is impossible to establish due to both philosophical/logical and empirical grounds, but that individuals do *perceive* and *experience* that which seems real. While we cannot "prove" a single objective reality, we can interpret and construct our subjective realities, as well as come to a consensus over what we in effect assume to be real (Polanyi, 1967). Scientific method is one method for coming to such a consensus, but I share with the constructivists the critique of "modernistic" science insofar as that science pretends to directly observe reality or establish the objective existence of the real.

3. In this thesis I distinguish between my view of rationality (e.g., the Free Will paper) and the critical perspective of rationality, common in the literature today, which characterizes "rational" models of behaviour as overly linear, deterministic, simplistic, and lacking human agency, values and subjectivity. This latter sense is always conveyed in this thesis by enclosing "rational" in quotation marks. Some, but not all, of such literature would share my view of rationality as reasoned, goal oriented behaviour, and accept the need for such a concept in the social sciences. Some of the literature, unfortunately, confuses bad models which pretend to model rational behaviour with the concept of rationality itself. Hence rejection of a rationality hypothesis on such a basis is, in my view, throwing out the baby with the bath-water, leaving us with *description* of human behaviour but little to *explain* it.

Contemporary scientific norms do not require that every researcher, and even less so, every research project or thesis, should clarify an ontology. Ontologies are largely implied by those parts of method which are explicit, by the way in which a researcher justifies her selection of research problem, and by the methodology chosen. Furthermore, whole disciplines, schools, or research milieux may share very similar ontologies; such fundamental belief systems become invisible to the insider as they become a matter of shared culture rather than individualistic philosophy.

This thesis does engage in an explicit discussion of ontology however, for two reasons. Firstly, as this is basically theoretical research which relies more upon the theoretically plausible conceptual connections among observations (which are always secondary) than upon directly tested correlations and cause-effect relations, there is less implied with respect to ontology by the methodology than is apparent from the majority of positivist research (i.e., which focuses primarily upon observation, whether quantitative or qualitative). Certain theoretical approaches will be favoured here partly because of the beginning ontological assumptions; it is better for readers to understand what those assumptions are so that they can decide whether they wish to reject arguments based on those assumptions, or follow my line of thinking either in sharing the assumptions or in remaining temporarily agnostic about them to see where they might lead to within the research problems posed here.

Values and Prescription. The second reason is that my approach to science is intimately intertwined with values, and *explicitly* so. My definition of science is a loose one, but it does contain limits: that science is a way of developing and transmitting knowledge whereby the knowledge produced has some independence from the methodology used (i.e., the process is not tautological or based entirely upon faith) and the methodology is to some degree reproducible (I say to some degree because I doubt that any but the most superficial techniques can be entirely reproducible). Few social scientists today will assert that any research practice is, or can be, value free. Many will assert that values inevitably affect, and should affect, the choice of research object, but that values affecting the way in which that object is viewed and interpreted should be identified *in order to* minimize their impact[4].

What motivates me to do science is a desire to play a part in improving the world — improved according to my values. This motivation helps to explain why I chose a research field which I call "policy science"; as I conceive this science it makes a direct connection from values — both individual and collective — to theory and data, not so much to describe reality, but to improve our experience of it. The fields of political science and sociology, for example, are more directed to *observation* of how policy is developed than toward *prescription*, which is rare in those fields and, when provided, is secondary and sometimes quite superficial[5].

However, to be consistent with my definition of science given above, in order to be a scientist who wishes to incorporate values in my research, , it is incumbent upon me to clarify what values are most important for driving my method, the implications for those values upon my ontological assumptions, and the criteria that will underlie the prescriptions that I develop and favour. Combined with epistemology and methodology, this allows for the development of knowledge which consists of policy prescriptions and associated analyses that provide a reasoned hope that they will advance the values that drive this research. I hope that the way in which such knowledge is developed could be adapted for use by others who share similar assumptions and values, or it least that the nature of the results could be evaluated with respect to how closely they honour them.

Rationality and Explanation. This part therefore incorporates a discussion of the ontological underpinnings of this research, in the manuscript "Free Will,

---

[4] In my opinion this is equally true of many research approaches which consider themselves "postmodern" alternatives to a modernistic science characterized as pretending that science is or can be a value-free endeavour. Many of them respond to the critique of modernism by shifting the object of study to discourse and symbols, but retain an equally positivistic approach to the study of those objects, while still aiming (and in some cases pretending) to be uninfected by the observer's subjectivity and values.

[5] Superficial in the sense that the links between what is observed, values and objectives, and policy prescriptions, are inadequately developed and argued.

Rationality, and Teleological Explanation: Toward a Teleological *Human* Science" (the "Free Will paper"). That paper starts out with saying that we *should* adopt the assumption that we have free will, for value-based reasons, and follows the implications of such an assumption for the nature of a *human* social science. The paper argues that a free will assumption implies rational goal-seeking behaviour on the part of individuals and collectivities and, furthermore, that the hypothesis of rationality is irrefutable.

Hence the role of the social sciences is to seek *rationality* in what they observe, and to help *advance* that rationality by reference to explicitly stated and developed values. I characterize such a science as *teleological*, in that its aim is prescription, given what *could* be, rather than mere observation.

Please refer now to the manuscript "Free Will, Rationality, and Explanation: Toward a Teleological Human Science".

I would like to convince my readers to share my ontological assumptions. However, for the purpose of this thesis what is most important is that they be understood. Once these assumptions are clarified, I hope that the nature of the ensuing research and the pertinence of the values advanced in that research will be also clearer. That research should be evaluated partly on the basis of its coherence with my ontology as I have described it.

Epistemology: A General System Theory (GST) Perspective of Complex, Dynamic Phenomena

Since the 1970s my approach to constructing, organizing and understanding phenomena has been imbued with systems thinking due especially to the book by Gregory Bateson (1972), *Steps to an Ecology of Mind*. Also, I have had long familiarity with basic GST concepts due to my background in the administrative and policy sciences, which rely heavily on concepts drawn from GST. Systems thinking fits very well with my general orientations having roots in political theory (Plato especially), philosophy (Hegel, Feuerbach, Fromm), and fiction (Steinbeck, Asimov, Shaw). These orientations portray humankind and the universe as intimately related in a dialectical relationship which can be viewed as teleological and inseparable from fundamental meanings of the "good".

Only in my doctoral program have I thoroughly investigated GST[6]; in so doing I found that it characterized fairly well my epistemological leanings which had previously been largely implicit. The role of GST for me has been to help structure inquiry and organize conceptually objects and phenomena in order to both recognize and simplify complexity, as well as to place objects in dynamic interaction over time and space. I share the view of the seminal writers of the field (von Bertalanffy, Ackoff, Bowler) that GST, as such, is an *epistemology*, not a methodology, and that it is devoid of values and empirical content. A GST *model* of reality requires the addition of empirical theory, and it may be made value-based by incorporating in it explicit value assumptions.

GST alone is a *formal* theory of what happens to "systems" under given conditions. In effect, it provides a language for thinking which reflects well how we already do think in the sciences, when we think systematically. We all *categorize* objects and phenomena, giving them labels and properties, placing them in groupings — often hierarchical, which is a tendency in systems models though not indispensable to GST approaches — and theorizing as to the implications of such groupings and phenomena as they work together over a period of time.

What GST suggests to the researcher is to be as explicit and self-conscious as possible about how objects and phenomena are and should be characterized and grouped, recognizing that *boundaries are always created by the researcher rather than given by reality.* Such thinking can facilitate the shifting of the researcher's frames of reference and encourage precise and systematic methods of structuring and interpreting reality. Awareness of how different *types* of systems interact, given assumptions, aid in theorizing dynamic behaviour over time — i.e., change processes which might also change the nature and boundaries of the interacting "objects".

My paper "General System Theory as a Postmodern Epistemology for the Social Sciences" (the "GST paper") is included in this thesis in order to describe GST as an epistemology and outline its development and influence in the social

---

[6] My examen de synthèse (comprehensive examination) was entitled "A Systems Theory Perspective of Mental Health Services and Policy Formation" (1994c).

sciences. Systems thinking has permeated all of the work in this thesis, although not all of the papers make explicit reference to GST. It is this approach which has encouraged me to take a very broad perspective of my "object of interest" both in terms of interacting and contextualizing systems and in terms of processes over time.

Please refer now to the manuscript "General System Theory as a Postmodern Epistemology for the Social Sciences".

Problématique: Reform Blockages in the Mental Health System

The mental health system has frequently been characterized by academic writers as a "system in crisis" (Goldie and Fredén, 1991; Kemp, 1991; Meyer, 1985; Prior, 1991); nevertheless, the media and politicians pay little attention to the problems of the system except on the rare occasions when the public's interest is piqued because of a dramatic event or poignant portrayal (Simmons, 1990; Warner, 1985) [e.g., the release of a movie about asylum conditions, an investigative series by a newspaper, deaths of itinerant people from exposure in front of the White House, a shocking crime or suicide].

Not surprisingly, given the lack of sustained public attention to fundamental mental health issues, analysts of mental health policy have frequently observed that major reform initiatives (e.g., community care) usually fall far from full implementation, and that the system remains virtually intractable with respect to meaningful reform (Bloche and Cournos, 1990; Kiesler, 1992; Marmor and Gill, 1989; Mechanic and Rochefort, 1990).

Major reform initiatives tend to end up pleasing neither reformers nor those favouring the status quo, since the actual results, after compromise, implementation, and adjustment, often appear to be inconsistent with any coherent mental health care paradigm (Jones, 1988; White and Mercier, 1991b). The most obvious example of this is "deinstitutionalization": a term which over the last twenty years has come to describe the most salient processes of the mental health system of the past 40 years (Bachrach, 1989).

Deinstitutionalization

Deinstitutionalization was accompanied by highly vaunted expectations that on the horizon was a new "biopsychosocial" mental health care paradigm, focusing on the patient as a whole person whose psychosocial as well as medical needs should by met by and within community (Kiesler and Sibulkin, 1987; Rochefort, 1984; Tyhurst, Chalke, Lawson, McNeel, Roberts, et al., 1963). However, despite the apparent initial support of major system actors (Grob, 1991; Hollingsworth, 1992), including the American Psychiatric Association (see Geller, 1994), the reform process became derailed and coopted (Becker, 1993; Goodwin, 1989).

Deinstitutionalization might be seen as the worst of both worlds, in that many individuals needing care, particularly psychosocial supports, neither receive it in institutions nor in the community (Beck and Parry, 1992; Boudreau, 1986; Isaac and Armat, 1991; Rachlin, 1989). Psychiatric care today, whether in institutions or "in the community", relies almost exclusively on the administration of psychoactive drugs (Harris, Hilton, and Rice, 1993; Kiesler and Sibulkin, 1987), which is not only inadequate, but often carries major risks of iatrogenic illness (Breggin, 1990; Cohen and McCubbin, 1990).

Many of the most seriously disabled are no longer in asylums but have been "transinstitutionalized" to other institutions (Brown, 1985; Morrissey, Goldman, and Klerman, 1985; Simmons, 1990) such as nursing homes (Shadish, 1984), jails (Warner, 1989), and private hospitals (Kiesler, 1992), or are receiving no care at all, being jobless and homeless[7] (Boffey, 1984). In the United States, while the mental health inpatient *proportion* of patient care episodes fell dramatically between 1955 and 1990, from 77% to 26%, the *number* of inpatient episodes per 100,000 population has continued to grow (from 795 to 917) (anon., 1994b).

Every indicator points to incessant growth in the mental health empire, particularly in the United States, with the targeting of new client categories, including children (Strauch, 1997) and those stressed by contemporary upheavals in the workplace (Kleinman and Cohen, 1991). "Analysis of the data obtained during National Depression Screening Day [U.S.] in October 1992 showed that 75 percent of those screened scored positive for depression and that 83 percent of

---

[7] It should be noted that a proportion of those "transinstitutionalized" should never have been users of the mental health system in the first place; e.g., the "mentally deficient" (Radford and Carter Park, 1993), and orphans such as the "enfants de Duplessis" (Cohen, 1992). Also, it might be speculated that a number of the extremely poor "deinstitutionalized mentally ill" living in the street were institutionalized in the first instance primarily due to conditions resulting from poverty and social isolation (Fox, 1978); the helplessness engendered by institutionalization could only worsen their incapacities once released into the community (Staples, 1993). The extent to which institutionalization, treatment, or being jobless and homeless is caused by or contributes to mental illness remains an open question (Kanter, 1989; Perlin, 1991).

those identified were not in treatment" (anon., 1994c). Rather than reform of the system, especially with respect to the most severely distressed, the foregoing figures combined with transinstitutionalization suggest rather a major *expansion* of the system.

System Outcomes

The basic challenges to and failures of the mental health system in the western world are long-standing, despite the various attempts at reform of and within the system (Kiesler, 1992). The system has changed its appearance in terms of institutional structures and laws, in terms of the language used in psychiatric diagnosis and etiological explanations, and in terms of the more visible characteristics of treatment.

However, there is general agreement among both critics and supporters of the system that its *outcomes* have not improved, and in several respects have worsened (Isaac and Armat, 1991). These and other commentators bemoan the increases in the numbers of those *considered* mentally ill, the numbers of the mentally ill who are homeless or jobless (Lamb, 1994), the magnitude of treatment-induced illnesses (Breggin, 1991), the unresponsiveness of the psychiatric system to patient problems, the infantilization of patients (Ingleby, 1985), the exacerbation of patient trauma (Jennings, 1994), and the extent of involuntary interventions (Szasz, 1994).

In this century the numbers of those labeled mentally ill have risen steadily; the prognosis for the more serious chronic illnesses has hardly changed — although neuroleptics can now mask some symptoms while creating others (Cohen, 1994a); and many individuals still have their liberties constrained in various ways, now including "compulsory community treatment orders" (Boudreau and Lambert, 1993a and 1993b). We seem to have made very little progress within the mental health system under most indicators — indeed there are strong indications of backsliding in recent years: Thompson (1994) suggests that cost-cutting since the 1980s has accompanied a new "dark age" in the mental health care system.

Perspectives of System Failure

The problem which this research addresses is, therefore, a mental health system which appears extremely weak in its ability to achieve substantial reform over the long run, either through normal administrative and scientific change (Regan, 1987) or through the occasional major concerted attempts at reform (e.g., the U.S. *Community Mental Health Centers Act*; see Rochefort, 1984). Various factors have been identified in the literature as hindering reforms aimed at improving outcomes, but there is no consensus on which factors are essential. This is not surprising since there are widely diverging ideas about what mental illness is and how society should deal with those considered afflicted (Boudreau, 1987).

Psychiatric Perspectives. For example, most psychiatrists today share a medical model orientation, part of a positivist view of science (Frankford, 1994), that leads them to view mental disorders as diseases having primarily genetic or biochemical origins and requiring the same types of treatment as other somatic conditions. Psychiatrists' dissatisfaction with the current system stems from their belief that large numbers of mentally ill are not having their needs met because they are not receiving treatment for their illness (Mercier, 1989), because they or their representatives claim excessive legal protections (Torrey, 1988), because the state interferes unduly in therapeutic practice (Lamb, 1994), and because hospitals are under-funded (Balter and Uhlenhuth, 1992). Most of these beliefs imply that what individuals need is not necessarily what they say they want.

User Perspectives. There is so little direct presentation of user views in the academic literature that the only generalization possible with respect to their views on reform blockage is that users themselves have little power within the policy system (Chamberlin, 1990). Among the relatively few users active in self-advocacy, two distinct perspectives have developed (Everett, 1994). The "consumers" accept the medical model but want increased choice and access to services (generally stressed in terms of private or public health insurance coverage). They hold that the major impediment to improvement is lack of public awareness about the prevalence and seriousness of mental illness, and of the progress medical science has been making and could continue to make given adequate and increased funding[8].

On the other hand, the "survivors" mount various critiques, some of them in the form of fundamental challenges to the purpose and constitution of the mental health system. They typically share concern about forced treatment and other constraints on the liberties of individuals. They see reform as impeded by other actors in the system with more power than users and with faulty ideas about user needs and about the importance to the therapeutic process of patient sovereignty (Salem, 1990; Oaks, 1993). Survivor groups regard biomedical psychiatric treatment with suspicion, as more or less objectifying and controlling users, and responding more to the interests of non-users than the needs of users. They tend to favour the application of a non-coercive model of community care and support, informed freedom of choice among a wide variety of alternatives — whether biomedical or psychosocial — and, more generally, a mental health system founded on empowerment principles.

One of the major impetuses to reform in the system, at least ostensibly, was concern over patient rights and liberties (Brown, 1984; Wald and Friedman, 1978). However, while many new procedures, rights and protections were passed into law, they have frequently turned out to be formalistic, with relatively little impact on the average patient considered for mandatory confinement or treatment (Holstein, 1993). Furthermore, most treatment is implicitly coercive rather than legally mandated (Lurigio and Lewis, 1989; Miedema, 1994; Reed and Lewis, 1990), especially where acceptance of "treatment" is a precondition to the provision of income (see anon., 1994a) or psychosocial services (Fineman, 1991), including income and housing.

The most fundamental problem, which renders finding a policy solution that much more difficult or impossible, is that we do not understand the major conditions classified by psychiatry as mental illnesses, such as those labeled as schizophrenia or bipolar disorder (Barham, 1986). Indeed, some critics have

---

[8] It is, however, difficult to separate out "consumer" views from non-consumers who advocate for them, since the "consumer" discourse is most frequently put forward by groups dominated by non-consumers (e.g., Ontario Friends of Schizophrenics, 1993). Those groups which restrict membership to consumers tend to adopt a "survivor" orientation.

suggested that what we call "mental illness" can be better understood as social or economic deviance. Such critics can point to the enforcement of treatment and other constraints to the liberties of those labeled mentally ill as only being understandable in an explanation about society's moralistic attitude to or fear of the mentally ill (Brown, 1985), considering the well-documented confusion of psychiatry in diagnosing and failure in curing emotional distress (Kirk and Kutchins, 1994; Mirowsky and Ross, 1989; Szasz, 1990; Wakefield, 1992).

Academic Analysis of System

In developing a narrative about the mental health system authors are hampered by the lack of agreement as to what constitutes the *needs* of the patients and therefore what a reform consists of. The literature can be seen as part of the system, reflecting the interests and ideologies of the writers (see MacIntyre, 1973; Segal, 1993). It is multidisciplinary, rather than interdisciplinary, spread out over various professional journals. With rare exceptions (see below), the overall system framework has not been directly addressed; actions are judged as good with reference to the disciplinary art (*lege artis*) rather than according to broader societal interests (*lege societatis*) [Lolas, 1994].

The result has been a focus on policy initiatives rather than on the context within which policy is developed, and on countless explanatory variables that do not fit together coherently in a way which would allow for explanation rather than description. This leaves implicit and unexamined the assumptions, raised as problematic in the historical treatments by Rothman (1980, 1993) and Simmons (1990) that (1) major reform can be achieved, (2) it can best be achieved incrementally within the system, and (3) on the whole the system operates *for* persons considered mentally ill. Wolin (1969) noted that the assumptions underlying a social system frequently remain unexamined:

> If society is conceived to be a system of decision-making, and if the recurrence of unjust decisions is commonly acknowledged, it follows that the system is, to some persistent degree, a structure of systematic injustice, otherwise the idea of a system is an inadequate account (p. 1065).

System-Level Analysis. The mental health sector is today almost always referred to as the mental health system. However, explicit use of systems thinking must lead to the recognition that any system is part of a larger system. Followed to its logical conclusion we are finally faced with fundamental ethical and metaphysical assumptions (Bowler, 1981). Hence, analysis of the mental health system as a systemic inquiry must be critical, asking whether failures are essentially internal technical malfunctions of a system that generally functions well, or rather basic outcomes of a system which might, depending on the value-laden criteria applied, be considered to be poorly constituted.

Social control literature. The most well-known systematic analyses of the mental health system have been developed by the "social control" (Brown, 1985; Castel, Castel, and Lovell, 1979; Foucault, 1962, 1972; Fox, 1978; Ingleby, 1985; Scheff, 1975; Scull, 1977, 1979) and the loosely related "anti-psychiatry" schools (Breggin, 1991; Hill, 1983; Szasz, 1974, 1991; see also Illich, 1977). Social control theorists use sociological or political economy paradigms (often influenced by Marxist thought) about the nature of society and the roles of institutions to argue propositions that the system is designed to meet broad societal/ideological/economic ends rather than to meet the needs of the mentally ill.

This literature is often profound, theoretically driven and extremely well-researched; it contributes substantively to understandings of why and how society has defined and marginalized the mentally ill. However, in order to have more impact upon policy development, the social control literature needs to recognize and animate human agency in order to break the "inevitability of failure" (Renaud, 1978) given our general cultural, economic and social environments of today.

Antipsychiatry literature. The anti-psychiatrists criticize the "medical model" paradigm, which they assert characterizes a mental health system which by virtue of a pseudo-scientific ideology is diverted from responding directly to the needs or desires of users. This literature contributes pieces of the puzzle to an understanding of the mental health system, but rather than systematically analyzing the system it tends to contrast characterizations of system phenomena with theoretical propositions regarding the nature of humanity, desirable values, or conceptions of mental illness (e.g., Laing, 1967). The degree to which readers have seriously considered this literature has too often depended, as with the social control literature, upon *a priori* acceptance of the often radical or eclectic conclusions of the authors. Those conclusions nevertheless rest upon rather ordinary assumptions about suffering, disease, etc.[9]

Both of these approaches have identified several reasons why persons labeled mentally ill suffer power disadvantages within the system, and why the system is not designed primarily to meet their needs. However, they apparently have had little impact on most of the recent mental health policy literature. It is unfortunate that the most radical of the analyses have been relegated to the fringe (Dain, 1989; Kenig, 1992), because regardless of whether one shares the world-view or models of mental illness of these critics, many of their observations, e.g., regarding conflicts of interest between those in the system with power and those without, could be understood and observed without a Marxist or "anti-psychiatric" theoretical stance, as was demonstrated by Goffman in his early (1961) sociological analysis of the situational dynamics of a mental patient in a psychiatric institution, and by Parsons (1974[1957]) in a "structural/functional" model of a mental hospital affected by supra- and sub- systems (see also Pilgrim, 1990). These analyses were important as being early applications of rudimentary systems thinking, creating a research agenda which with rare exceptions has not been taken up[10].

The Policy Playing Field

What the literature requires is a "bridge" which enables entry of critical approaches but does not depend upon any particular highly controversial presuppositions, or upon disciplinary approaches inaccessible or unacceptable to other disciplines, or inapplicable to the policy level. A bridge might be

---

[9] I am grateful to D. Cohen for pointing this out.

[10] In recent years rigorously systematic theory driven research seems to have fallen out of favour in the human sciences, in favour of either modernistic positivism (which may rely on qualitative as well as quantitative methods) or "postmodern" relativism (theory free analysis based on discourse rather than value-defined social problems; this includes some variants of social constructionism [e.g., Spector and Kitsuse, 1977]).

constructed by evaluating the policy "playing field": even if system participants disagree as to what patients "need" and how delivery and administration of health care should be structured, are there not some *process* standards of "fairness" or "efficiency" that we could substantially agree should apply with respect to formation of policy? Is it possible that regardless of what we think is wrong with the mental health delivery system we might be able to agree that the mental health policy system is bound to diverge widely from meeting the "needs" of persons diagnosed as mentally ill — *however* we may conceive those needs?

The Needs-Based Rationality of the Mental Health System. The literature has never addressed, systematically, the theory and assumptions, explicit and implicit, that would underlie a belief that the mental health system is meeting or can meet the "needs" of users. This is not to say that the foundations of system rationality based on meeting needs have not been shaken, e.g., by debate over etiology, therapeutic orientation, and the ethics of coercive treatment. But these debates need to be taken together and viewed from a higher systemic level. For example, aside from the views of particular individuals or professions as to what patient needs are, and how they should be met, does the mental health system as a *policy and politics system* have the capacity and the will to rationally moderate those debates and tend toward the "better" answers based on reason, argument, science, and fundamental human values?

The article "Deinstitutionalization: The Illusion of Disillusion" (the "Deinstitutionalization paper") examines what has happened to the major reform thrusts of the last 50 years in the mental health system. It shows that despite a wide initial consensus over the importance of psychosocial supports in the community, progress has been disappointingly slow in implementing that consensus — indeed, that what started as a movement out of asylums and into normalizing community care and support has, to the extent that there has been "deinstitutionalization" rather than transinstitutionalization, been a movement more toward coercive chemical and crisis care than community care, received by individuals isolated in the community and periodically hospitalized.

Please refer now to the article "Deinstitutionalization: The Illusion of Disillusion".

Did the form that such "reform" took represent social choice, rationally based on some theory of how user needs should be determined and assessed? The deinstitutionalization paper and this thesis have continually raised that question, leading to the suggestion that the workings of the system can be much better explained by reference to the *interest-based strategies of those in the system with power* than by a narrative about the system functioning to meet the "needs" of users. The question will then arise as to whether there is any reason to believe that such self-interested strategies will result, in any event, in user needs being met.

Research Objectives:
Toward a Critical Theory of the Mental Health System

The objective of this research was to contribute to the development of a critical theory explaining why the mental health system has been so immune to meaningful reform. The research aimed to create a "road map" of the system in terms of the most fundamental categories of variables that could facilitate or impede reform. Therefore it needs to have a broad enough scope to incorporate variables that might be excluded in analysis of a narrow definition of a mental health system. Interactions involving agents and institutions need to be understood not only in terms of immediate causal links but also in terms of long term patterns and strategies.

While the principal results of this particular research are not intended to be prescriptive, the research is undertaken in order to facilitate the development of subsequent policy planning or other prescriptive work[11]. However, the research should not limit itself to questions which would be apparently the most fruitful or feasible with respect to immediate policy prescription, since it is already apparent that the system is beyond "easy" solutions.

A theoretical proposition is described in the next part; it is designed to provide a systematic way to structure the research around fundamental characteristics of the mental health sector as a dynamic system: this proposition was chosen for its fecundity in terms of developing sub-hypotheses, for its capacity to direct the researcher's attention to questions that might not be addressed in more narrowly defined or strictly positivistic research, and for its robustness in terms of its ability to order new information and reorder old. Research in testing and elaborating the theoretical proposition, both in this thesis and beyond, should therefore advance our knowledge of the mental health system.

This statement of objectives focuses on the ability of the theoretical proposition to stimulate research partly because the thesis had been conceived to

---

[11] This thesis is intended to lay the ground-work for a career research program in mental health policy science, which aims to develop policy orientations and prescriptions firmly grounded in a critical theory of the mental health system.

be a "thesis by articles". It was never expected that any individual article, or all of them together, would fully test the overall proposition, although those used to assess the proposition will be shown to be derived from it. While the research and articles generated are wide-ranging in terms of topic and disciplinary fields, they are not eclectic, since each provides pieces of the puzzle: that puzzle being the inability of the mental health system to achieve long-lasting meaningful reform.

From the inception of this research it was expected that the benefits resulting from the research conducted pursuant to the theoretical and analytical framework presented here would be, even if the findings were to turn out to be inconclusive:

- To obtain and transmit a more thoroughly systematic critical understanding of the mental health system as a *system*: i.e., as a whole in interaction with its parts and with supra-systems.
- To raise the level of analysis of the "mental health system" — i.e., as a *social* system with political and economic dimensions, not just as a "delivery system" scientifically constructed by ideology-free experts to meet the objectively determined "needs" of users.
- To partially explain system mutations and characteristics teleologically, in terms of *human agency*: strategic decisions of humans, individually and collectively, within social, political and economic contexts and processes.
- To understand the difficulty the system has in developing *"rational" policy*, or at least policy that has a reasonable likelihood of satisfying objectives concerning the well-being of users.
- In the spirit of the interdisciplinary classic by Robert Dahl (1961) analyzing the use of power (including economic resources) and the mechanics of influence in the "political system" of an American city, to develop an analytical framework which will facilitate policy development in the mental health policy field and provide a *policy science* approach which could be adapted to other policy fields.

Dissemination of Results and Expected Benefits

Within the broader aegis of critical political economy approaches in policy science, this research has aimed to develop a General System Theory analytical framework which draws heavily from the field of public choice, applied to mental health policy but adaptable to other areas of social policy.

As a "thesis by article", at least three articles ensuing from the research project must have been published or submitted for publication. This thesis incorporates three published articles, two submitted articles, and four manuscripts in advanced form (all of those are being disseminated as *GRASP Working Papers*). The four manuscripts and an article based on the overall results contained in this thesis will be submitted for publication after revisions. Hence this thesis will have generated up to 10 articles.

Although each article will reach a somewhat different readership, due to the nature of the articles and the journals, in general it is expected that results of the research will reach at least four primary audiences:

- Persons with substantial involvement in the *mental health policy* field, including academics in various social science disciplines, and policy planners/analysts in governments and related environments.
- Those *critics of mental health* practice and policy who are interested in a broad, integrative, interdisciplinary approach to understanding the social and policy contexts and interrelationships of mental health issues.
- Specialists in "*policy science*": an approach to policy analysis which is inter- or adisciplinary (or a discipline in itself), which draws on theory and methodology from fields including political science, economics, sociology and geography. This new field has been more heavily influenced by systems theory than any other and can be expected to welcome a contribution which develops a systems theoretic framework for use in policy analysis, and which has demonstrated its applicability to a particular major area of social policy.
- Persons interested in the intersection between *ethics*, social as well as clinical, and psychiatry and mental health practices in general and, more precisely, those interested in the ethical and other philosophical issues involved in mental health law and policies.

I hope to use the work completed in this research as the basis for a policy sciences textbook, which would demonstrate an integrative, systematic, systemic, interdisciplinary approach to modeling a policy field, using the mental health field as an example or case study.

Theory and Research Hypotheses: Conditions for the Proposition that the Mental Health System Meets User Needs

As discussed above, few commentators are satisfied with the mental health system; it has been variously described as being permanently in crisis, irrational, doomed to failure, incapable of meaningful reform, oppressing those it pretends to help. If true, such characterizations indicate fundamental systemic weaknesses: that the system does not adequately achieve its purpose, insofar as that purpose is to relieve the distress of those considered to be mentally ill and help them to function with dignity within society.

The Proposition that User Needs are Met

The specific proposition to be critically assessed in this research is the following: *the mental health system, as presently structured, can be expected to satisfy user needs*. This is the alternative to the null hypothesis that patient needs are not met. The term "expected" is key to the type of research proposed here. Rather than attempting to directly test, by operationalized empirical measures, the "degree" to which patient "needs" are met, the approach used here seeks to ascertain whether there are *reasonable theoretical grounds* for such an expectation. If there are, we are justified in rejecting the null hypothesis that patient needs are not met.

Falsifiability. By directing this research toward finding reasonable grounds for a proposition that user needs are met I am therefore seeking grounds for rejecting the null hypothesis that user needs are not met. Since a belief — at least from a scientific or policy perspective — that user needs are met *requires* something like "reasonable grounds", the *null hypothesis is refutable* — i.e., if the null hypothesis were not true there would be some way for so ascertaining (see Popper, 1979).

In order to falsify a hypothesis that patient needs *are* met we would have to do one of the following:

1) Define "needs" *a priori*, operationalize them empirically, and verify the degree to which they are met. This is, of course, problematic, insofar as "needs" can be

defined by different actors in ways that suit them — which has been, in effect, the usual case.

2) Set up conditions which, if met, create the *expectation* that needs will be met, *even among people sharing different conceptions of need.* Falsifiability of a hypothesis that needs are met requires that it is possible to demonstrate that the conditions are *not* met if in fact they are not.

Selection of Null Hypothesis. Hence, not only is my null hypothesis falsifiable, but so would an alternative when posed in the sense of 2) above (but not as easily, as will be discussed below). This thesis does set up what I believe to be such conditions. Hence it is almost arbitrary as to whether the hypothesis that patient needs are met should be considered the null hypothesis or the alternative to the null. The decision is not totally arbitrary, however, due to considerations of scientific conservatism and ease of refutability.

Conservatism. The null hypothesis is conventionally considered the conservative position, with which we rest unless there is convincing evidence otherwise (e.g., $\partial < .05$ in statistical hypothesis testing). There are several candidate criteria for determining which position is considered conservative; two of the most common are the prevailing beliefs and consequence of error (whether it would be more costly to wrongly reject or wrongly accept a hypothesis).

I accept the latter view, which reflects both a critical approach to science — the implications of which include healthy skepticism with respect to "received wisdom" and "common sense", understanding that such forms of knowledge are often ideologies constructed to justify or obscure systems of privilege — and a policy science inclination which requires that justification of government programs, particularly costly and coercive ones, should take non-efficacy as the conservative position — the burden of proof lies with showing that such programs actually work.

While many would agree with this attitude *in general*, it seems to me that the *mental health system* relies much more upon *faith* than proved usefulness, for a variety of reasons which this thesis partly addresses. From both policy science and critical theory perspectives, therefore, it is inappropriate to select the null hypothesis solely on the basis of the status quo or common beliefs.

Ease of refutability. In general, it is easier to establish the existence of something, if it exists, than the non-existence of something, if it does not exist — i.e., *non-existence is more refutable than existence*[12]. This issue was discussed in the Free Will paper, where I pointed out the irrefutability of a hypothesis of rationality. Establishing that something is being done to meet needs might, similarly, be easier to establish (if it is true) than the converse.

Therefore, I designate in this research the hypothesis that patient needs are not met as the null hypothesis, and assess reasons and evidence to the contrary.

Decision Choices. *A priori*, this research can therefore lead to one of the two following results:

1) *Acceptance of the hypothesis that patient needs are met.* This requires convincing evidence of meeting a sufficient condition. For the result to be credible, so must be the character of the condition as sufficient.
2) *Non-acceptance of the hypothesis that patient needs are met.* To be credible, this requires enumeration of all those sufficient conditions that might reasonably underlie a theory that patient needs are met, and convincing argument showing that there is no good reason to think that any of them are met. Note again that this outcome does not mean *rejection* of the hypothesis, merely failure to accept it.

In addition to one of the above outcomes, it *might* be possible to both *reject* the hypothesis that patient needs are met and *accept* the null hypothesis that patient needs are *not* met. This would require that the hypothesis that patient needs is met is *fully* refutable — i.e., that there is good evidence that *no* sufficient condition is met, and that there are reasonable grounds to think that there is such a thing as "user needs" and that such a concept can be defined as an empirical object.

---

[12] Take, for example, the question of the existence of God. While we cannot "prove" the *nonexistence* of God — at least not without defining *a priori* what God is supposed to be and having conditions for such existence which can be falsified if such a God does not exist — we can critically assess various theories and narratives which ostensibly establish the existence of God. If we find them wanting, we have not proved that God exists, nor have we come to a "definition" of God. We have just failed to accept the proposition that God exists.

This is because if the latter were not possible, it would be impossible to enumerate all potential sufficient conditions and hence refute them.

This will not be an alternative in this thesis, because I do not intend to introduce and defend a specific conception of needs that has empirical content. Indeed, I am not sure that the concept has usefulness with respect to policy analysis, independent of other concepts such as preferences, rights and values.

Therefore, **it is important to emphasize that the way the research problem is structured in this thesis does *not* require an *a priori* definition of needs, nor, with respect to choices #1 and #2 above, even acceptance that there *are* user needs.** Rather, explicit and implicit narratives about how they might be met — which do not always have to depend on a clear definition of needs — are described or constructed, and then assessed.

Sufficient Conditions for a Hypothesis of User-Centered Rationality

This method of argumentation then proceeds by identifying those major arguments which might underlie confidence in the system with respect to its *user-centered rationality.* By "user-centered rationality" I mean an explicit justification for the mental health system as an expression of social choice (the creation or maintenance of a valued good by a collectivity such as a nation) whose criterion is the welfare of users, however defined[13].

Pursuant to the teleological ontology and GST epistemology which guide my approach to understanding social phenomena, it is a simple matter to identify sufficient conditions for the expectation that the system will meet user needs, by assuming a broad mental health system whose activities result from individual and group strategic decisions made in complex, uncertain contexts over time. These contexts, in turn, are structured by socio-economic patterns of interaction and exchange and constrained by distributions of resources such as money,

---

[13] This paper assesses the rationality of the system only with respect to the criterion that it meet the needs of users. Hence, failure to find systemic rationality under that criterion would not exclude other grounds, such as protecting the public, maintaining cultural or economic norms, or providing employment. To evaluate the system properly, it is necessary to *conceptually* separate the different objectives of the system.

influence, location — that is, *power* (Black, 1986; Crozier and Friedberg, 1977; Etzioni, 1982; Scott, 1985).

I have constructed below three conditions any of which, to the degree that they are met, would support the hypothesis that the mental health system will meet user needs. Each incorporates a class of narratives which have been or could be advanced as supporting the hypothesis. Taken alone each statement, if true, *might* provide a sufficient, but not *a priori* necessary, support for the hypothesis. These statements are:

1) **Interest convergence condition**: *Powerful actors in the system have interests which converge with those of users.*
2) **Power access condition**: *The system satisfies user needs as a result of user power to affect system outcomes — which power might be exercised in various ways throughout the system (at the political, policy, administrative, and/or therapeutic levels).*
3) **Policy intervention condition**: *Insofar as the system does not meet user needs — e.g., insofar as the interest convergence and power access conditions are inadequate — policy interventions at the administrative and service delivery levels can compensate.*

The Interest Convergence Condition (*The Paternalistic Model*). Powerful actors in the system have interests which converge with those of users. This means that sets of interactions in which actors pursue their own interests — however different from those of users — are likely to serve the interests of users. This condition does not *require* that patient "needs" are those that they express, but that powerful actors are truly motivated to intentionally or unintentionally meet "true" patient needs whether or not those needs conflict with what patients say. This condition represents, in effect, the paternalistic narrative, as will become clearer in the part discussing the results of the thesis.

The Power Access Condition (*The Political Model*). The system satisfies user needs as a result of user power to affect system outcomes — which power might be exercised in various ways throughout the system (at the political, policy, administrative, and/or therapeutic levels).

This condition implies that outcomes desired by users are those that they "need" — if not for every individual then on the whole for users as aggregated or acting collectively. This then conflicts to some extent with the interest convergence criterion, in terms of how needs are determined. This condition may be seen as consistent with implicit or explicit conceptions of "needs" as *desires* or *claims;* such conceptions are found more or less in various competitive models that focus on power as a key explanatory variable. They might be grouped into two classes: the "survival of the fittest" models which are essentially descriptive but possibly containing a naturalistic ethic (that what is good is that which has the strength to survive, defeat competitors, and reproduce), and the "fair competition" models based on theories of liberty, autonomy, capitalism, democracy, and postmodern (or identity) politics. The fair competition models assume or require some kind of egalitarianism in the means of access to power, and place confidence in such an egalitarian "market-place" for efficient or fair allocation of values among those competing in it.

The Policy Intervention Condition (*The Bureaucratic Model*). Insofar as the system does not meet user needs — e.g., insofar as the paternalistic and political models are inadequate — policy interventions at the administrative and service delivery levels can compensate. Hence this condition views collective or State action as correcting for failure to meet needs because of a political "market-place" which disfavours users and a failure of relatively informal or intermediate power structures (caregivers, family members, corporations, community groups) to protect the interests of vulnerable users.

It is important to note that this condition is not independent of the interest convergence and power access conditions, insofar as policy interventions affect the interests and powers of actors and especially — as will be further elaborated — insofar as powerful actors in the system control the nature of policy interventions.

Also, while establishing whether the first two conditions are met does not necessitate an evaluation of the degree to which "needs" *can* be assessed and met — the interest convergence condition assumes that parties sharing interests also share needs and the power access condition assumes that what is sought is what is needed — this condition includes situations where needs have to be *defined,* objectively *assessed,* and where needs deficits can be *measured.* This would *not* be the case insofar as policy intervention is aimed at improving the degree to which the first two conditions are satisfied rather than toward government processes which themselves try to assess and meet needs.

Theoretical Interpretation of Findings

Satisfaction of any of the above conditions is not likely to be completely dichotomous (fully satisfied vs. complete failure to satisfy). Therefore, unless all three propositions are explicitly addressed together we have not fully evaluated the user-centered rationality of the system. For example, one might assert that interests of the various categories of system actors largely diverge, but the benefits of the system are allocated fairly due to relatively egalitarian access to power, or because of the laws we have established regulating medical delivery systems.

The very expression of the latter position of course raises the question as to *how we would expect laws to protect patient interests would come to pass in a system where users have little power and those who do cannot be expected to act in accordance with user interests;* such a question might not be thought of under a different theoretical approach. Alternatively, one might posit that while users obviously suffer severe power disabilities, it is reasonable to believe that their interests are complementary to those of caregivers.

However, strong arguments tending to rejection of all three propositions/conditions would seem to leave us with little choice but to rest with the null hypothesis; i.e., that there are no grounds to expect that the mental health system meets the needs of users. If we do remain with the null hypothesis then we have failed to establish the rationality of the mental health system insofar as it is based upon meeting user needs. Such a conclusion could be argued against by advancing and supporting some other sufficient condition than has been outlined here. I would welcome such argument as part of dialectic in the research community. However, several years of research have failed to reveal to me such a candidate.

If the research were to reveal reasonable support for one or a combination of the conditions outlined here, then at least a *prima facie* case will have been made that the mental health system meets user needs. Critics of such a conclusion might then wish to examine the assumptions contained in the supported conditions. For example, is it really true that those whose interests are convergent with users can be expected to satisfy user needs as a result of self-interested actions? If users controlled mental health policy, would they use that power to satisfy their "true" needs?

However, regardless of the degree to which the outlined conditions are really sufficient conditions for meeting user needs, those conditions capture, in my view, the predominant narratives giving rise to such expectation. Hence, if none of them are met, we would be left without any reason to think that user needs are met. Given the evidence presented earlier in this thesis regarding reform blockages, such a finding would be very disturbing indeed, and would suggest the urgency for a very thorough-going rethinking of the fundamental assumptions of the system and how it is constructed.

Methodology: Toward a Mental Health Policy Science

This part will, firstly, describe the methodological orientations of the thesis as a work of policy science, drawing upon political economy and founded upon a teleological philosophy and GST epistemology. I will then describe particular analytical methodologies used to achieve these orientations, which pertain to public choice theory, rational choice and other issues of decision theory. Thirdly, I will describe how the research process was carried out and how information was obtained. Finally, given these methodological aspects of the research, I will then discuss issues of internal and external validity.

Disciplinary Orientation: Interdisciplinary, Systems-Theoretic Policy Science

This research is consistent with a career research orientation based upon a functional goal rather than an object or traditional discipline: to contribute to the making of better social policy. To this end, I have obtained training in political science (specializing in public policy and administration) and economics (specializing in microeconomics, welfare economics, and public choice theory).

However, the new interdiscipline of "policy sciences" captures best my disciplinary orientation. Interdisciplinary by definition, it is centered around the process of policy formation and, as opposed to political science, is primarily normative and applied, rather than descriptive, in the sense that it is problem-oriented and seeks to improve policy outcomes or processes rather than merely understanding them (DeLeon, 1988). Many policy scientists draw heavily on the research found in other disciplines to this end.

The thesis research incorporates pertinent literature from psychology (e.g., community or social psychology), sociology (which has provided much of the extant analysis of the mental health sector as a social system), economics (to understand pertinent market based activities, to enhance evaluation analysis, and to facilitate analysis of choice processes), philosophy and ethics, social work, therapeutic and health related disciplines (e.g., nursing, epidemiology, public health, medical geography), and medical professions (since these disciplines or professions are part of the object of study, but also because they contribute to policy analysis and evaluation).

The new discipline of policy sciences, like GST, was born in the 1950s and developed rapidly in the 1960s. Like GST, part of its *raison d'être* was its inter- or multidisciplinarity as a response to the perceived failures of the insular sciences to adequately address human problems in a world growing rapidly more complex and threatening. Indeed, to varying degrees over time the systems and policy science streams virtually converged: "systems analysis and policy analysis are used as essentially synonymous terms for the same activity" (DeLeon, 1988, p. 24, citing Majone and Wade). However, while the fundamental defining characteristics of the policy science discipline are interdisciplinarity, problem orientation, and contextual placement of problems in a policy *process* (DeLeon, 1988) — all of which are apparently *consistent* with GST — it would be incorrect to view it as the application of GST to policy problems.

The policy science field arose out of the realization that problem solutions no longer became evident merely by studying the problem: required instead was a generic expertise in teasing out the hidden, and recognition of the crucial importance of policy development, implementation, evaluation and adjustment as negotiated processes over time, in competition with other problems, under constraints and uncertainty (see Bosso, 1994; Rochefort and Cobb, 1994).

Limitations of Policy Science. However, this ideological justification for the field was not fully met in practice. While the policy sciences could have profited from GST philosophy as it developed since the 1960s, stressing conceptions of open dynamic non-linear systems moved by values, goals and interaction, in the early years of the field policy scientists became to a large extent policy scienticists, selecting from the systems field *tools* which made their work seem scientific, methodologically rigorous, cutting edge, etc. In their appropriation of techniques of operations research, systems analysis, cybernetic concepts, linear programming, decision and game theory, etc., they merely *extended* Taylorist methods, in order to better manage humans and information, rather than adopting a new paradigm more appropriate to the second industrial revolution.

These early policy scientists had too great a faith in the ability of technique to deliver "rational" results. This was partly the fault of those who "stripped down" theoretical models, ignoring the often *explicit* caveats and limitations accompanying those models. Popular books and other media made fashions of various techniques. Since expectations for newly fashionable techniques were unrealistically high, their life-spans were short.

After frequently noted repeated failures in policy development, implementation, and evaluation[14], the postmodern critique, and the recent interest in non-linear dynamics and process orientations, linear quantitative or technicist research is no longer uncritically accepted. However, to the extent that the backlash against "rational" policy analysis consists of theory-less relativism (postmodernism at its worst), or structure-less reductionism (chaos theory at its worst), the field will not provide fertile ground for reform.

Perhaps a field such as policy science will always oscillate between emphasis on analysis and synthesis in accordance with larger societal, policy and methodological trends, but there is a range beyond which the field can become so unbalanced that we may not know what to make of what we observe, or do not know how to make the transition from theory to practice. This thesis suggests that systems thinking can help maintain such balance, insofar as it stresses process *and* structure, deduction *and* induction, analysis *in* synthesis, and, particularly, the ability of human systems to change themselves.

Mental Health Policy Analysis. Mental health policy analysis at the system level is dominated by the new policy scientists — particularly those with backgrounds in political science (e.g., David Rochefort: Rochefort, 1993; Rochefort and Portz, 1993) and sociology (e.g., Deena White: White, 1992; White, 1993; White and Mercier, 1991a and 1991b). The sociological literature dealing with mental health policy, while often lacking explicit theory, has contributed valuable and refreshing conceptual approaches derived from the theoretical traditions of sociology. Some of the policy scientists and sociologists have been influenced by systems thinking or use system tools (e.g., Hastings, 1986; Hollingsworth, 1992), which should be

---

[14] Policy science evaluations, appearing in the 1970s and 1980s, often repudiated the great policy initiatives of the 1960s and 1970s. However, the validity or meaningfulness of the evaluations has also been questioned.

expected since the object of study is typically referred to by policy analysts as "the mental health system".

David Rochefort's prolific body of work since the early 1980s, consistently of high quality, reflects well current mainstream mental health policy analysis as a "policy science". Rochefort implies in "Approaching Mental Health Policy Analysis" that he follows an "eclectic" approach, using a variety of methodologies; in this it is influenced by Wildavsky as combining "art, craft, and science" (Rochefort, 1993, pp. 8-9). Rochefort and colleagues address systemic issues like coordination, development, and efficacy, sometimes using systems concepts that have become commonplace in the policy sciences (feedback, cycles, resources, strategy, system hierarchy, information needs, ecosystem, system fragmentation, "dynamics"). By design, however, Rochefort's work lacks a consistent theoretical or analytical framework:

> Two maxims accompany a methodological eclecticism in policy analysis that stresses policy content over analytic procedure ... First, no single methodology can capture the full variety of data that are of interest. Second, the aspect of the policy topic being worked on should determine choice of methodology, not the other way around (1993, p. 10).

The above refers to methodology; there is no similar discussion with respect to theory, "eclectic" or otherwise. While a great deal of the theory-free descriptive literature in the mental health policy journals is quite rich, pondering many important factors and issues within and around the system, the repeated failures in social policy reform in general, and in mental health policy in particular, suggests that there is a need for more integrative work — not simply by adding more variables and dimensions but by applying a method of analysis appropriate for incorporating complexity and dynamism within a coherent theoretical framework which reduces complexity while increasing coherence and comprehensibility.

Lack of *a priori* theory, and of a methodological approach pushing the researcher "from behind", inevitably results not in an *inquiring system* — a concept used by Shakun (1981a, 1981b) in modeling the GST approach to policy analysis — but what I would term as an *attracted system* whereby the researcher is drawn to the available data and unconsciously adopts well-established perceptions. While breakthrough thinking need not be the result of "a complete contempt for historical facts" (Jones, 1988, p. 86, referring to Foucault), it is unlikely to result from merely gathering and reporting the facts, unless that process is inspired by a vision that exerts itself in determining what types of facts will be generated or gathered, and in how they will be organized and interpreted.

This argument is made in my paper "Mental Health Policy from a Systemic Perspective: Orientations for Strategic Reform" (the "Strategic paper"); it shows how policy analysis in the mental health field has failed, both in terms of method and results, describes what a systemic perspective implies for mental health policy sciences and, consequently, for the nature of reform strategies. In effect, it proposes a methodology for mental health policy science which represents the essence of the methodology behind most of the research in this thesis.

Please refer now to "Mental Health Policy from a Systemic Perspective: Orientations for Strategic Reform".

Toward a Theory-Grounded Model of the Mental Health System. The research in this thesis has progressed toward the creation of a model of the mental health policy and politics system by progressively adding characteristics about the "system" to be explained (assumptions, statements, observations, and hypotheses). A number of these characteristics have generic types, in that they have been discussed or defined for species of abstract systems by GST theorists (structures, processes, degree of openness, complexity). Assumption of system "types", in conjunction with other theory, speculation, and observation, implies expected system *processes*.

This results in, then, a theoretical framework within which a series of hypotheses can be deduced, some of which are directly operationalizable and hence testable in order to support or not the seeming veracity of the hypotheses and theory. The term "framework" is used intentionally, denoting a model with sufficient clarity in its structure and components that it can be improved, critiqued and verified, and compared with other relatively clearly defined models in the literature.

In the end, the aim has been to design a model such that, along with the application of value-based stances, the formulation of strategic planning for reforms of the mental health system will have something approaching a scientific basis. Clearly among the prime requisites for the efficacy of such a model is not only its ability to model a heterostatically dynamic system, but to be so itself. Hence it needs to be robustly built, more like a Leggo set than as a puzzle; i.e., a change in one piece should not cause the whole edifice to fail.

Analytical Methodologies: Public Choice and Rational Choice

Many articles in the thesis, and especially the overall problématique and theoretical structure, are influenced by public choice theory and perspectives. The field of public choice theory is interdisciplinary, straddling political science and economics, and drawing on theories or observations from many human and natural sciences, especially anthropology, sociology and biology. Mueller (1989, p. 1) defined it as "the economic study of nonmarket decision making, or simply the application of economics to political science". Much of the thesis analysis can be characterized as "public choice" by virtue of its application of microeconomic analysis to mental health policy and politics[15].

Rational Choice. The fundamental assumption of public choice is rational choice, which in my applications is based on the utility theory developed in microeconomics. This assumption holds that individuals are rational, in that they attempt to maximize utility, or satisfaction, given the circumstances they are in. Such individuals, given their basic values, make decisions, often strategic, in an environment of incentives, constraints and uncertainty. The rational choice approach has been most developed on a theoretical basis within the fields of microeconomics and game theoretic decision theory and, with respect to public goods and collective decisions, in the related area of public choice theory (see Cornes and Sandler, 1986; Katz, Nitzan and Rosenberg, 1990; Mueller, 1989).

As opposed to mainstream health economics, which focuses on questions of health delivery efficiency, rational choice approaches have only recently been

---

[15] E.g.: short term "selfish" strategies as a "prisoners' dilemma" problem (Deinstitutionalization paper); comparison of the relative effectiveness of lobby groups with respect to the "free rider" problem (Unbalanced paper); application of power (political science) and expected utility (microeconomics) to substituted decision making (a non-market choice process) [Error paper].

applied to *mental* health policy (e.g., Gigliotti and Rubin, 1991). Explicit assumption of rational behaviour provides a means for analyzing and predicting how power and interests are translated into action, and how constraints and incentives impact upon the choices of actors within the system (see Bates, 1993).

In the case of those diagnosed as mentally ill the assumption of rationality is problematic, since coercive or paternalistic behaviour toward persons so labeled is partly based on the assumption of irrationality or incompetence of mental patients (Appelbaum and Schwartz, 1992; Rose, 1986). In research problems where the assumption of rationality is problematic it has been implicitly treated in this thesis as a proposition, alternative to the null hypothesis of irrationality (as explained in the Free Will paper). One test for such a proposition is whether "irrational" behaviour can be understood as rational given the individual's circumstances (basic interests or values, power, incentives and constraints), thereby allowing rejection of the null hypothesis.

Utility Theory. The Error paper relies explicitly and extensively upon applications in decision theory of utility theory under conditions of uncertainty. Utility theory has been criticized on a number of grounds touching on questions of hedonism, materialism, lack of quantifiability, problems in interpersonal comparison, inability to serve as a value-free technical tool in the policy sciences, theoretical tautology, and inapplicability to human behaviour that is irrational, deontological or altruistic (Etzioni, 1988).

Mainstream economics has virtually abandoned utility theory, replacing it with methodological tools such as "revealed preference" (Houthakker, 1950), "rational expectations" (Muth, 1961; Nerlove, 1958) and "Pareto optimality". These changes have resulted in the displacement of economics from the human sciences to the area of technique (closely related to — or a sub-field — of mathematics and statistics) [Paldam, 1993; Parker 1993], insofar as an emphasis on description and macro statistics based prediction of choices has displaced concerns about understanding them.

While all of the critiques directed at utility theory or its applications have some or much merit, most of them depend on rather narrow conceptions of utility and its application. Utility theory — even when it is understood that "utility" *per* *se* does not explain the fundamental values underlying choice — is useful in the policy sciences for two main reasons:

1) In choice problems it focuses attention on how choices are made rather than what the choices are; emphasis on the latter can neither explain nor predict (except in the sense of probabilistic predictions based on past patterns of choices by aggregates of individuals), merely describe. Emphasis on the former can sometimes help to explain the latter.
2) Policy options that incorporate a recognition of utility distributions, although the quantities of such utilities cannot be measured or precisely estimated, place decision making responsibilities in the hands of policy makers rather than technicians. The policy makers, e.g., the voting public and elected officials, in effect subjectively estimate gains and losses from policy alternatives and include those estimates among the decision making criteria. The resulting decision is frequently justified in accordance with these estimates.

The approaches of mainstream economics tend to ignore that which is not quantifiable and use a "Pareto optimal" criterion. A situation is Pareto optimal if no party can improve without another party becoming worse off (Varian, 1984). While Pareto optimality is a useful criterion for *efficiency* — in short-term analyses — it has nothing to say about either fairness or social welfare. Because it takes initial distributions of wealth and resources as given, the extensive reliance upon it by policy makers using economic analysis has had the effect of strongly favouring the status quo both theoretically and practically, especially in matters of complex social policy (Samuels, 1989).

Human nature. Ironically, some non-economic fields, e.g., sociology, have also rejected utility theory, ostensibly because of the associations of that approach with conceptions of human nature as selfish, hedonistic, materialistic, etc. (Caillé, 1986). Despite the fact that its originator, Jeremy Bentham, did entertain such conceptions, many of such critiques are naive, because utility theory *per se* contains *no* assumptions about human nature and what is or should be valued, other than the basic assumption of rationality: that people will try to get what it is that they want, given constraints and opportunity costs. The fact that utility theory developed mostly within the field of economics influenced the way it was used and coloured the common impressions of it. However, some economists, especially within public choice, have shown how utility theory can be consistent with a large variety of ideas of humans, e.g., as altruistic, masochistic, deontological (Buchanan, 1972).

Irrationality. Finally, observed "irrational" decision making behaviour has been held to weaken utility arguments. It appears however that such "irrational" decisions can often be seen as rational within a utility approach when a meta-level of analysis is invoked within a systemic perspective (e.g., larger time frame; set of, rather than discrete, decisions; contextual variables). The proposition of rationality in the face of apparent irrationality has often led to greater understanding of the decision making process, of the actor, and of the actor's values (see Pettit, 1993, pp. 264-283) [for example, analysis of psychiatric patient "non-compliance" with medication].

Decision Theory and Public Choice from a GST Perspective. The boundaries among the decision/game theory, system, and policy science fields have overlapped since the contributions of Herbert Simon (e.g., 1969[1956]), who is a founder of rational choice theory, a systems theorist, and an eminence in administrative science. The systematic approach of decision/game theory in modeling decision/negotiation/power relations in situations of uncertainty, constraint, and complexity, suits the theoretical conceptions of GST and the detailed applications interests of systems analysis and operations research.

Groff observed in 1976 that organizational decision making still relied on classical analytical models, failing to utilize basic systems concepts. Recent GST approaches to decision making emphasize certain aspects that had been lacking in early formal "rational" approaches. The most important recognition was that *goals are not given*; they develop. Individuals' values and goals change as they and their environment change, and group goals are negotiated, often in situations of power and conflict (Shakun, 1981a; Tanaka, 1989).

Secondly, problems are not defined *a priori*; they are *subjective* cognitive frameworks (Bullock III and Rodgers Jr., 1976; von Foerster, 1988) with fuzzy borders in a *system* of problems (McNeil, 1991). Third, decisions are not discrete but rather *development processes*. As processes they can be modeled as systems incorporating feedback, adjustment, negotiation, and *alteration of the decision system* *itself* (Contractor and Seibold, 1993; Kuroda, 1993; Rosen, 1974). Apparently "irrational" decisions, contraventions of economic theory, paradoxes, etc., might be understood when it is realized that, as described in the Error paper, decisions are *strategic*, having consideration of a variety of uncertain scenarios *over time* (Bierman, Jr., 1989; Grossman and Watt, 1992). "The attempt to deal with a system of problems as a system — synthetically, as a whole — is an essential property of *planning* in contrast to problem solving" (Ackoff, 1974, p. 5).

Finally, decisions are frequently — particularly and increasingly in the public policy sphere — complex, non-linear, dynamic *social* processes, in which individual choices or stances at one point in time depend on the perceived choices or stances of others (Latané, Nowak, and Lieu, 1994). Outcomes might not be understandable without understanding the process dynamics (well described by Hofferbert and Schäfer, 1982), which may be chaotic (De Greene, 1990; Kiel, 1992; Richards, 1990). Not only are *goals* value-based, but so might be the decision making *style* (Snow and Bloom, 1992).

According to Ackoff (1974), policy scientists should be less concerned with *what* decisions are made, but *how* they should be made. Such concerns have always been evident in the work of Amitai Etzioni, who has played a major role in the development of both decision theory and the policy sciences[16]. Etzioni used a systems approach to address the problem of fraud and abuse of patients in nursing homes (1982). The policy sciences have suffered in recent years because of spectacular failures of major policy initiatives. However, Etzioni suggests:

> What to the uninitiated are unanticipated consequences of initiatives are typically the work of other factors that have been overlooked by tunnel vision but that have a chance of being seen by the systems analyst (p. 36).

Etzioni used *constituency analysis* to "identify the constellations of social and political power that sustain existing opportunities for fraud and abuse, and those

---

[16] The work of Etzioni takes on different significances in different fields, since he has very wide interests and has made important contributions to a variety of research areas. He has also, for example, contributed ethics and value theory to microeconomics (e.g., 1988), and has been a leader in the recent development of communitarian theory (e.g., 1995) — both contributions having influenced this thesis.

that will support countermeasures" (p. 26). With respect to social power, his proposals related to prevention rather than the inefficacious punishment/deterrence model. With respect to political power, he delineated the interests various classes of people can be expected to serve, and concluded that families of elders in nursing homes "seem to be insufficiently motivated to act on their behalf" (p. 30); hence he proposed incorporating groups like the Gray Panthers in the reform process.

Etzioni's more fundamental policy proposals were rejected[17]; his frustrating experience with attempts at reform illustrates the paradox of the policy sciences: "reform" in the *interior* of the system may be easy to instigate but unlikely to succeed; reform of the *system* might succeed, if only it could be implemented. Depending on the circumstances, a reform strategy might justifiably be internal and incremental or external and revolutionary[18].

The above discussion provides methodological and analytical principles regarding choice, decision making, rationality, and policy strategy in contexts structured by values, uncertainty, negotiation, constraints and incentives, and power. These principles provide the essential concepts drawn upon in this thesis for understanding outcomes of the mental health system that might otherwise be obscured by ideological narratives which are themselves system manifestations, or which might be incomprehensible in an analysis which is overly simplistic, static, and lacking faith in human agency as rational, purposive behaviour in complex environments. Without such a more profound and systematic understanding of the mental health system as propelled by individual and collective choices, it

---

17. The governor's appointed commission rejected Etzioni's idea for creation of a permanent commission involving persons with a commitment or incentive to patient interests, holding its own hearings instead. In the end, ten "innocuous" commission recommendations were adopted; the only one not adopted would have prevented members of the legislature from receiving money from the nursing home industry! It would be interesting to follow up this issue, to see to what extent the adopted "reforms" succeeded in reducing fraud and abuse over the long-run.

18. Or, given disillusionment or powerful interests conflicting with reform objectives, either simply giving up (e.g., income security policy) or providing the pretense at reform by periodically changing labels and administrative structures (e.g., penal policy).

would be — as discussed in the problématique, including the Deinstitutionalization paper — impossible to identify the "policy levers" that need to be pulled in order to disengage the enduring policy gridlock that has marked mental health policy[19].

Research Procedure

The thesis is primarily theoretical, reflecting a research process described above as an "inquiring system", whereby the empirical information utilized has been sought pursuant to research questions deductively derived from the theoretical structure. Therefore, while the thesis can be regarded as a form of "research synthesis", in that it draws from already available research rather than generating new data (as opposed to information), it is certainly not a "meta-analysis" (Windle, 1994) which, properly speaking, is a systematic analysis of analyses — often purely descriptive — rather than a transformation of existing research within a fresh theoretical framework independent of that research.

Hence, the research relied heavily on the use of reason and logic in a synthetic dialectical process with available data, analysis and critiques. For each hypothesis, an initial understanding was developed of the nature of the system — including sub-systems, system contexts, actors, and interactions. Given this understanding, an initial analytical framework was developed in which terms were defined (e.g., "rationality" in the Free Will and Error papers), necessary and sufficient conditions identified, falsifiability criteria determined, and a "map" created which categorizes and enumerates concepts, including their linkages and overlaps (e.g., the characterizations of system actors and their interactions, in the Deinstitutionalization paper). This process resulted in a number of problems or

---

[19] As well as, it might be added, most if not all other important areas of social, economic and environmental policy; this thesis does not pretend that the larger questions with respect to achieving major enduring reform in the mental health sector are substantially different in kind from other major policy sectors. All of them could, I think be analyzed in ways similar to the approaches guiding this thesis as complex systems characterized by what will be seen to be extreme divergences among actors in terms of power and interests — which divergences are often obscured and distorted by self-interested ideological narratives.

propositions to be addressed through subsequent analysis, or study of the literature or available data.

For example, what kind of process would be required in order to determine, with reasonable validity and reliability, the "best interests" of a user considered medically incompetent (the Pure Best Interest and Error papers)? Once such conditions are outlined, failure to satisfy them would suggest that the "best interests" of a patient are not met where decision making is substituted. While the thesis usually provides some indication of the degree to which such conditions are satisfied, using argument and drawing upon empirical observations contained in the literature, its main contribution lies in, I hope, the development of "better" questions (Bachelard, 1969).

Answers to some of them may become very apparent once the question is posed — e.g., do severely distressed users have well-organized and well-financed lobby groups[20]? Others may be sought by other researchers, or by myself in the future — e.g., what decision making process do legally mandated substituted decision makers actually follow in making treatment decisions on behalf of patients considered incompetent?

Research proceeded in focused directions within each of the articles. The answers to specific questions or discussions of specific issues were sought within the published literature. This focused research has required continual revision of the analytical framework[21]: concepts redefined, new propositions or logical structures developed, etc. New versions of the reworked analytical framework often in turn pointed to further research. Hence this research process involves iterative steps in terms of shifting attention between system levels (e.g., user

---

[20] It cannot be sufficiently stressed how important it is to this research, as to critical theory research in general, *to state the obvious* — and not only to subsequently question it, because too often banal and obvious verities become virtually invisible to the point that researchers fail to follow up on their causes and implications.

[21] E.g., earlier versions of the thesis proposal stated the overall proposition as "the system can *not* be expected to meet the needs of clients ..."; this was changed, since the proposition that the system *can* be expected to meet user needs is much more falsifiable — i.e., by demonstrating that conditions for such an expectation are not met.

"needs" at the macro-system level, dealt with in the Unbalanced paper, and at the micro level, dealt with in the Pure Best Interests paper), and between theory and data[22].

It is important to note that it is rarely possible to provide in a published article, which must be short and to the point, a complete or explicit description of the research program which gave rise to the article, including its overall theory, epistemological underpinnings, and analytical method. Also, the articles do not all present themselves as applications of systems theory, even though the "reality", abstractions, and logical arguments are organized according to the systems thinking presented in this thesis.

It should be expected that articles prepared for publication in a particular journal with a target audience will have been influenced by the necessity for compromise, both before submission and after peer review. The articles themselves could be considered as sub-systems of the thesis: the articles need to be interpreted in light of the context provided by the thesis, in order to be understood at the thesis level. Accordingly, the thesis "context" — the text that accompanies the articles — describes in further detail how each article was generated from the overall research program.

Information Sources. The thesis research did not involve the formal generation of new data not already available in reports and journals, although existing data has often been processed or reinterpreted to meet the needs of the project (e.g., explanation in the Deinstitutionalization paper of how changing contexts changed the strategies of psychiatrists, affecting their support for [and definition of] "community care"). The research was developed within a teleological philosophy and GST epistemology, within a policy science orientation, and created a systemic

---

[22] For example, there was substantial evolution in the development of the Deinstitutionalization paper with respect to how the role of "economic factors" was interpreted within a causal explanation of deinstitutionalization (for an earlier version, see 1994a). Initial hypotheses generated literature searches which resulted in restatement and refinement of the hypotheses, which, in their new form required further research, in an iterative process which required deadlines to come to an end.

analytical framework, drawing upon political economy including public choice: all of these fields are interdisciplinary by definition.

Therefore, in addition to public documents and the journals of several human science disciplines having pertinence to mental health and the mental health system (political science, psychiatry, psychology, sociology, social work), sources included literature in a variety of fields providing philosophical concepts and analytical and methodological tools (economics, decision and game theory, communications). Reference was also made to natural sciences such as physics and biology in obtaining insights into complex or chaotic dynamic systems (see, e.g., the GST paper).

Hence a large body of multi-disciplinary literature was used in this research, generated as follows:

- Approximately 30 of the most pertinent journals (health, systems, public choice, policy science) have been checked periodically since 1990 and screened systematically for articles of potential value to my research objects or analytical approach[23].
- Bibliographic databases were searched for literature regarding major objects/topics of interest[24] as well as for analytical approaches applied to other

---

[23] These journals, spread out over 7 libraries at Université de Montréal and McGill University, include: *Hastings Center Report; Social Science and Medicine; Journal of Health and Social Behavior; Journal of Health Economics; Milbank Quarterly; International Journal of Law and Psychiatry; Santé mentale au Québec; Policy Sciences; Public Choice; Review of Social Economy; Journal of Health Politics, Policy and Law; Policy Studies Journal; Psychiatric Services; Community Mental Health Journal; Journal of Mental Health Administration; New Directions for Mental Health Services; Journal of Medicine and Philosophy; Culture, Medicine and Psychiatry; Canadian Journal of Community Mental Health; Administration and Policy in Mental Health; Canada's Mental Health; Systems Research; Behavioral Science; Journal of Psychiatry and Law; Journal of Mind and Behavior; Mental and Physical Disability Law Reporter; Bulletin of the American Academy of Psychiatry and Law; Issues in Law and Medicine; Health and Canadian Society.*

[24] Basically the same set of keywords, with some periodic adjustments and additions, have been used since the inception of my Ph.D. program to maintain my literature up to date — e.g., "(mental or psych$) and (system$ or polic$ or politic$ or power$ or empower$)". Since 1995 I have

objects/topics but which might be adapted to my research (e.g., with respect to preference estimation, for the Error paper).

- Citation searches were made with respect to particularly pertinent authors or papers (e.g., citations to Scott, 1985, as virtually the only paper which explicitly modeled as a system a broadly defined mental health sector).
- Books are not directly included in the commonly used citation databases, except PsychInfo, however book reviews are, and hence pertinent books were revealed by key-word searches. Important new books in my areas of interest are also revealed by book reviews in the journals periodically screened. Unfortunately, however, I was only able to draw upon a fraction of recent books pertinent to the thesis, since they are rarely available at the Université de Montréal library (I have had somewhat more success at McGill, and buying more than a few books is out of the question for most Ph.D. students[25]). To the extent that such books are scientifically important, however, they are usually based on articles already published in peer reviewed journals.
- Older literature, and pertinent literature not revealed due to the above search procedures, were often revealed as citations in the literature I had previously obtained.
- Books and articles have also been suggested or made available by the thesis director and others. This source of information has been particularly important with respect to recent books critical of current psychiatric practice, because they are less likely to be based upon articles in health journals, and are rarely available at Montreal libraries.
- Copies of pertinent literature, including virtually all articles I cite, are kept in subject files. I have developed a very extensive collection of literature aimed at serving the needs not only of the thesis but also my career research program.

---

relied primarily upon Current Contents to identify new literature. I frequently also use Medline, PsychInfo, Sociofile, and LegalTrac to identify older literature in specific areas.

[25] I think it is quite pertinent to point out here that while it is expected that Ph.D. theses be truly at the cutting edge, most will fall somewhat short due to resource limitations which affect students much more than established researchers. In my case certain resources were available to me due to my presence at GRASP throughout my program that greatly facilitated my access to pertinent journal articles.

Given that the vast majority of available literature is in the English language, because I am a Québecois attending a francophone university I have paid particular attention to literature from Québec or in the French language.

External and Internal Validity

The locale of the researched system, in general, is North America and Western Europe. While across those continents mental health systems are diverse, in important respects there are enough elements in common to allow for a reasonable degree of generalization and abstraction (e.g., the organization of the psychiatric and other helping professions, treatment contexts faced by patients, concerns of family members, formal political institutions).

The development of an abstract model or portrayal of the mental health system has drawn from a larger amount of literature and other sources of information from Canada and the United States, and to a lesser extent from some European countries, especially England, Scandinavia and France. Hence its applicability to particular sub-national or non North American countries will be limited or extremely limited[26]. However, since this research has tried to make its

---

[26] Countries whose language of academic expression is neither English nor French are underrepresented in this research, since I read only English or French. It is important to note, however, that increasingly, more important research from non-English countries is published in English. Nevertheless, this does create important gaps in the research. For example, I know very little about the mental health systems, and especially their political aspects, for Italy, Germany, and Spain. This lack is aggravated by the fact that *critical* literature has fewer international publication outlets in English. However, Current Contents does publish English abstracts for all major academic journals whatever the language, which has helped to reduce the ethnocentric bias of this thesis.

theoretical bases and empirical assumptions explicit, its application to particular jurisdictions or temporal periods could be enhanced by adjusting its parameters[27].

Since this research has not generated new data, its internal validity will be limited insofar as the availability and reliability of data for some questions is less than desirable. It was anticipated, pursuant to analysis of the theory driving this research, that certain types of information would be particularly sparse: the market for research is influenced by the existing distribution of power (political, social, economic) and interests. It is obvious, for example, that there is very little information in the academic journals regarding the values, objectives, and decision making processes of psychiatric or mental health users, particularly where supported by direct empirical study. The resulting problems of internal validity should not be too threatening to the results *insofar* as:

1) it is reasonable to believe that much of the missing information would, if it were available, tend to support the findings of the research, hence its absence skews results on the conservative side;
2) important information shortfalls have been identified and explicitly pointed out;
3) the research is heavily theory-driven such that to a large extent the type of information required is specified by the analysis, rather than the propositions being derived from the available information as an "attracted system" of research.

The thesis, and its arguments, should be seen as contributing to a debate: a dialectic within the research community whereby the participants construct, deconstruct, extend, and otherwise critique each other's arguments. This research "constructs" arguments that I see as *implicitly* supporting the assumption of user-centered rationality of the mental health system, and subsequently critiques those arguments. I will welcome criticism of the perspectives developed in the thesis as furthering the development of a systematic theory "explaining" the mental health system. Such criticism might question the completeness of the hypotheses or their conceptual clarity, or might present arguments countering the counter-propositions developed in the thesis.

---

[27] As opposed to positivist research, in which the results apply only to the population, as limited by time and space, from which the sample is drawn. Without theory to explain the observations, there can be no basis for generalizing or adapting the results to other populations.

Results: Evaluation of the Interest Convergence and Power Access Conditions

Analytical Requirements for Evaluation of Research Hypothesis

Dependence/Independence of Conditions. This thesis set up three conditions that, if met, would support the research hypothesis that patient needs are met and thereby permit rejection of the null hypothesis that they are not. Each condition, if adequately met, is sufficient for that purpose. This thesis operates on the assumption that those three conditions are exhaustive: i.e., to my knowledge no other hypothesis, explanation or narrative, with any degree of *a priori* plausibility, has been advanced in the literature that would support the research hypothesis. If the thesis concludes that we have no reasonable grounds to believe that the mental health system meets user needs, among the ways such a result could be criticized would be by a counter-argument which would seem to reject the null hypothesis on grounds other than the three conditions I had set up.

Note that each condition could be viewed as independent from each other in logical and statistical terms. In the logical sense, what is signified by the definition of one condition does not determine the signification of another. In the statistical sense, empirical variation in the degree to which one condition is met does not explain all variation in another condition. Hence, insofar as these conditions are independent in the senses indicated, evaluation of the proposition that user needs are met requires separate evaluation of each of the conditions.

However, there is no implication in this thesis that the conditions are *completely* independent. While they may be in a purely abstract logical sense, in that empirical circumstances might be imaginable whereby all conditions are free to vary without at all being affected by the others, it seems reasonable to expect some covariation in the real world — as with virtually any two variables defining phenomena over time and space.

Furthermore, we might have reason to believe that even if we were unable to justify rejection of the null hypothesis on the basis of evaluation of the degree to which each of the three conditions, considered separately in turn, are satisfied, one might argue that some lesser degree to which each of the conditions are satisfied will suffice to reject the null hypothesis, once the *cumulative* impacts of the conditions are evaluated. E.g., we might be able to argue that while users have little power and caregivers have little interest in meeting patient needs, there is just enough interest convergence and user influence for user needs to be fairly met.

This part will separately evaluate the first two conditions regarding interest convergence and power access. It will be demonstrated in the conclusion that the third condition, regarding policy intervention, is in fact quite dependent on the first two. Hence the conclusion will discuss whether reforms are possible, and what kinds of reforms, given the nature of interest convergence and power disparities discussed in this part. The conclusion, therefore, will discuss the ways in which the conditions might interact, and will evaluate their cumulative impact — given the possible, observed or likely interactions *among* the conditions.

The Unbalanced paper. The article "Extremely Unbalanced: Interest Divergence and Power Disparities Between Clients and Psychiatry" (the "Unbalanced paper"), written with David Cohen, is the key paper in the thesis, as it concisely presents and evaluates the theory in the thesis. It also provides a *problématique,* as well as directions for reform of the mental health system taking into account the discussion contained in that paper regarding the narratives supporting an expectation that user needs are met. As such, that paper represents a summary of the thesis. Given that the Unbalanced paper addresses all of the conditions separately and together, this should be read prior to embarking on the discussion below, which draws upon that and other papers in summarizing the thesis results for each of the conditions and finally for all of the conditions taken together.

Please refer now to the article "Extremely Unbalanced: Interest Divergence and Power Disparities Between Clients and Psychiatry".

Most discussion in this part centers upon the interest convergence condition, because it is analytically more complex but especially because there is such a preponderance of evidence tending to rejection of the power access condition that discussion of it amounts to, firstly, remarking on the obvious, and secondly, developing theory that explains the finding that user power deficits are both enormous and systematic.

The Interest Convergence Condition *(The Paternalistic Model)*

Assessment of this condition was structured within a logical framework constructed in the same way as for the construction of the three overall conditions — i.e., by seeking sufficient conditions for support of the narrative. Hence, on what basis might we think that the interests of users coincide with those of other actors in the system with power?

The Unbalanced paper described three narratives which would, to the extent that they were empirically well-founded, support a hypothesis of interest convergence: we characterized these narratives as the medical, market and best interests models. I believe that these models subsume the explanations that explicitly or implicitly have been advanced or which have underlain an assumption of interest convergence; if so, we can assume them to be exhaustive.

Medical Model. There are two major approaches to assessing the "medical model" as a means for satisfying user needs. One approach is to deny the existence or importance of somatic factors in the etiology of mental illness. This approach has gained little influence, probably because one cannot *prove* the *non*-existence of somatic causality. Furthermore, in the current context where medical model explanations of mental illness are pervasive to the point of hegemony, the burden of demonstrating evidence is borne by those challenging the medical model. Those questioning the somatic etiology of "mental illness" tend to concentrate, therefore, on showing the existence of non-somatic factors, or demonstrating that evidence for somatic causality cannot bear scientific scrutiny, resting rather upon fanciful wishful thinking.

The more influential critiques, notable in the work of Szasz, for example, suggest that the issue of actual etiology is rather beside the point, when the question is not one of etiology but of intervention — particularly paternalistic intervention. These arguments suggest that if somatic factors underlying disorders of *social and interpersonal behaviour* in psychiatric patients are identified (which they have not been), these would no more justify a medical approach to managing these behaviours than a medical approach is justifiable to manage the social and interpersonal behaviour of demonstrably heart-impaired or lung-impaired patients[28].

Consistent with the latter argument I will argue below, in effect, that the issue of whether a medical model can be used as a narrative suggesting that user needs are met does not require a belief one way or another as to whether somatic factors are present in the etiology of psychological distress. I will pursue this argument at two levels of analysis: the *micro-level*, pertaining to the doctor-patient relationship, and the *macro-level*, pertaining to appropriate policy directions with respect to services and research concerned with the health of *populations*. The conclusions of both discussions are, I believe, robust with respect to the possible role of somatic factors in psychological distress[29].

Micro-level: the doctor-patient relationship. Note that I have characterized the interest convergence condition as "the paternalistic model". The reasons for such characterization will become more apparent as this condition is analyzed below: it will be seen that the various narratives which rest mostly upon interest convergence — as opposed to another condition — are all paternalistic. In other words, they place confidence in those with power or socially approved authority to decide what user needs are and how they should be satisfied. The Unbalanced paper suggests that *even if* serious psychological distress is somatically based, the

---

28. My thanks to David Cohen for helping to clarify for me the subtleties in arguments counter to the medical model.

29. Although they may seem to lose some force to the extent that somatic factors are very determinative. In my view, a critique of my arguments based on such a position would be impossible to justify on the basis of systematic, replicated scientific evidence — there simply is not such evidence. A believer of this proposition would be of course entitled to such belief, but it would remain to argue why a public mental health system in a democratic polity should be based upon it. A meritocratic polity, assigning merit on metaphysical rather than scientific, democratic or other grounds which can be rationally justified, might permit the translation of such belief into public policy. Hence Szasz's frequent characterization of psychiatry as the "new priesthood" and psychiatric ideology as the "new religion". This thesis does not aim to convince biological "true believers" of the points of view taken here.

questionable scientific basis for diagnosis leaves open the use of mental illness labeling based upon deviance from social norms, influenced by socioeconomic, cultural, race, and gender differences between practitioners and patients.

Furthermore, again allowing for somatic bases for mental illnesses, that paper points out that such somatic bases do not directly point to the "right" treatment: every treatment alternative holds advantages and disadvantages for patients, which may be valued quite differently by patients as opposed to doctors. For example, reducing those symptoms which are considered more "medical" — or problem-causing — may reduce patients' quality of life by also reducing alertness, affect, and cognitive abilities. Also, every somatic treatment has possible iatrogenic outcomes, which can be severely disabling, permanent, or fatal (e.g., tardive dyskinesia; see Cohen and McCubbin, 1990).

A purely medical model approach to treatment is incapable of incorporating patient valuations and preferences with respect to treatment alternatives — which might include psychotherapy, community care, counseling, or even no treatment at all — which bear on patients' own life experiences, activities, cultural norms and expectations.

Rather, a strict medical model approach suggests *somatic* treatments — e.g., electroshock and/or drugs — for what are hypothesized to be strictly *somatic* conditions. Insofar as a mental health system based upon the provision of medical treatment for psychological distress responds to user preferences based on their *own* fully informed choices, and provides other alternatives for users whose self-assessed needs go beyond the medical, there would remain a reasonable narrative to suggest that user needs are met. However, the justification for such system is not derived from the medical model but rather from a narrative favouring consumer sovereignty.

But to the extent that the system is entirely medical, does not inform users of treatment risks and alternatives, and in many cases coerces or forces medical treatments, the relationship between such a system and meeting user needs becomes very tenuous. Such a system must be based on a paternalistic conception of user needs that regards determination of them as a purely medical question — as a medical question there would be no room for user perspectives or evaluations of alternatives. In the medical care system generally there are few, if any, defenders of such a proposition today; it is now well-recognized that both medical expertise and patient preferences are required to achieve optimal health care decision making.

Nevertheless, in the field of psychiatric care, the unquestioned authority of medical psychiatric care remains far more than a vestige, in both practice and in public attitudes. This attitude remains despite the recognition, when the issue is raised, that pursuant to the general acceptance of the principal of informed consent in health care generally, psychiatric patients should also be allowed to play the role of informed consumers, *unless* there is a formal determination of incompetence.

We are left, then, with no acceptable role for the "medical model" *as a narrative justifying a belief that user needs are met.* Appropriate application of medical treatment in the mental health system might be justified on other grounds, e.g., user choice (individually pursuant to a market model, or collectively pursuant to the power access condition, both to be discussed below), or by assessment of needs on behalf of *incompetent* users (e.g., under a "best interests" model, to be discussed below).

Otherwise, we are left with a justification that few would accept today, *if indeed the question is explicitly posed* (the rationale for this thesis is that such questions are not systematically posed and thoroughly addressed), that even with respect to competent psychiatric patients, authority to decide treatment should rest entirely with medical practitioners, due to their expertise.

I suggest that it is incumbent upon defenders of such a position to stipulate whether it should also apply to any competent patients facing the possibility of any medical treatment, and if not, why not. I am aware of no argument which justifies unquestioned medical authority on behalf of psychiatric patients, distinguished from medical patients in general, that does not revolve around, explicitly or implicitly, concepts of incompetence. A justification for authoritative treatment based upon patient incompetence is not, however, derived from the medical model, even if the causes of such incompetence are properly the subject of medical intervention.

The Unbalanced paper summarizes what are already well-known critiques of a medical model approach to determining and satisfying the "needs" of psychiatric patients. In the case of psychiatric care, the most important need under the medical model is "treatment" for an "illness". By definition, insofar as a *medical* paradigm is invoked to justify what happens to users in the mental health system, that paradigm rests on an epistemology which views or defines psychological distress or deviance as *symptoms* of a *physical disease.* By physical disease I mean, as the common understanding of the term implies, an etiology based *primarily* on somatic origin: e.g., genetic "defect", physical trauma, infection, usually implicating the brain.

The Unbalanced paper briefly mentions a failure to validate somatic pathology in mental illnesses, but does not rest its critique of the medical model as a justification for an expectation that user needs are met upon this issue. It should be noted, however, that if it is true that there is little or no evidence for somatic bases, as *primary* causal factors in mental illnesses, there would be *no basis for a medical model argument that a mental health system centered around psychiatric care — as a specialty of medical care — meets user needs.* This is not to say that such a system could not be justified, but rather that it could not be justified on the basis of a *medical model* argument (such a system might be justified, for example, on the basis of user *preferences* — a justification which does not rest on the actual scientific status of somatic etiology).

Most of the fundamental critiques of the mental health system are based primarily upon the validity of the medical model with respect to somatic pathology for psychological distress (see, e.g., Cohen, 1990 and 1994c). Such critiques are well-known to user activists, but are very rarely referred to in the psychiatric, mental health policy and administration, or even community care literature, even though many of those critiques are extremely thorough and even conservative with respect to the scientific evidence and arguments drawn upon to support their positions.

I am not qualified to fully assess the somatic etiology claims of psychiatrists who assert that mental illnesses are "brain diseases". I have noted, however, that such claims are usually *implicit* or *assumed* — the vast majority of psychiatrists, and mental health service systems centered around medical approaches, do not even question or evaluate such somatic bases. *Direct* evidence of somatic bases, e.g., genetic markers or brain lesions, is widely publicized when initially published, but is rarely, if ever, consistently replicated (Horgan, 1993).

Indirect evidence, e.g., efficacy of somatic treatments upon conditions hypothesized to be somatic in origin, has been very problematic on several counts.

Among the arguments more fatal for the credibility of such evidence is the non-specificity of treatments (treatments are not "magic bullets", in that they affect a wide variety of behaviours not specific to the hypothesized illness), the marginal efficacy of treatments compared to placebos, particularly *active* placebos, the curious tendency of treatments to lose much of their efficacy when compared to "new and improved" treatments (suggesting, it would seem, "test effects" including systematic investigator bias favouring the "new" treatments), and the failure of somatic treatments to significantly and consistently surpass non-somatic care (psychotherapy, lay counseling, community care) with respect to efficacy (although well-controlled longitudinal studies in this regard are rare — particularly efficacy measured in terms of subjective quality of life measures and functional measures regarding normalization and social integration, as opposed to simply a reduction in those symptoms defined in terms of psychiatric pathology). These issues are summarized in the review by Cohen (1994b).

Macro-level: the health of populations. This thesis takes a relatively agnostic position with respect to the existence or not of somatic causality in the development of psychological distress, thereby avoiding much of a debate which at least in recent years seems to have little impact upon mental health policy. Indeed, the directions that this thesis points to would be consistent with an etiology containing somatic elements.

However, consistent with the GST epistemology underlying this thesis, such etiology must be understood as non-linear, probably multi-factorial, and, related to those aspects, *interactive* among "causal" factors.

Furthermore, the teleological ontology which suffuses this thesis suggests that *deterministic* explanations of psychological distress, based on *physical* factors divorced from human choices, are inadequate. These points are made in the manuscript "Population Health: A Call for Breadth (Mental Health) and Depth (Psychosocial Theory)" [the "Pop-Health paper"], which discusses the implications of a *biopsychosocial* model of health and illness — which may incorporate "biological" causal factors — for the health of populations.

Please refer now to the manuscript "Population Health: A Call for Breadth (Mental Health) and Depth (Psychosocial Theory)"

The Pop-Health paper suggests directions for theory development which could combine both so-called "physical" health issues with mental health issues within a population health approach. That paper draws upon literature which shows the importance of *social* factors — e.g., control or sense of control, education, and socioeconomic status — for health outcomes. Such literature typically based its findings upon ill-health outcomes which had previously been closely associated with *biological* etiology, e.g., cardiovascular disease. Such literature rejects a medical model orientation for improving the health of *populations*.

This view suggests that a macro-level view of health systems renders questionable reliance upon *medical* approaches, recognizing that while medical treatment may be viewed as necessary by individual practitioners in individual cases where disease has already manifested, there may be, when considered at the collective level, more efficacious means of treatment and amelioration of disease and, especially, of ill-health *prevention*. Such a macro-level orientation permits the introduction of psychosocial *and* biological (e.g., air quality) causal factors in models of health outcomes that are virtually invisible at the individual level, or are ignored as beyond the reach of individual practitioners treating individual persons already sick.

Insofar as the role of somatic factors in the etiology of psychological distress is less than for "physical" diseases having established somatic pathology, and insofar as patient choices and attitudes can reduce the illness or negative impacts associated with psychological distress, compared with physical illness, the adequacy of the medical model for improving the health of populations would be even weaker in the case of mental health.

Furthermore, as described in the Pop-Health paper (in the section "the construction of etiologies of emotional distress"), the medical model orientation of current psychiatric research is poorly equipped to isolate not only psychosocial but also *biological* etiological factors. That is because the medical model orientation focuses attention on *ill*-health outcomes rather than upon processes of divergence from good-health or normality. Construction of an etiology under the medical model then becomes a matter of attempting to trace a path *backwards in time*; as discussed in the Strategic paper with respect to explanations of mental health policy outcomes, such explanatory methods suffer from hindsight bias, over-linearity, and over-reliance upon more visible and immediate factors.

Hence, while a somatic treatment might be favoured by individual users to reduce some symptoms of psychological distress that they view as undesirable, and might even be effective in some cases in eliminating the postulated defect (although I am aware of no such somatic treatment), persons suffering psychological distress, as a *collectivity*, might reasonably wish to see policy and research shifting emphasis toward the variety of factors — social, economic, environmental — which create conditions favouring the development of psychological distress. Such a perspective would recognize that even if somatic "determinants" are important in "causing" psychological distress, those determinants are also influenced by non-somatic factors.

The population health approach recognizes that the question of *intervention*, with respect to reducing ill-health outcomes for entire populations over time, is quite a separate question from that of "*causality*". For example, even if a genetic predisposition is present in all cases of "schizophrenia", and were more determinative in such an outcome than any other single factor, it is nevertheless not obvious that we can more effectively intervene in populations at the genetic level, compared with intervention at the level of social policy. In the policy sciences, the concept of "policy lever" recognizes that *the point of effective policy intervention is not always the point where we would identify the "key" causal variable in terms of strength or immediacy.*

The Pop-Health paper therefore advances a systemic "biopsychosocial model" for health and ill-health outcomes at the population level; this model provides a narrative which can explain all that the medical model does but with much more depth over time, space, and populations, and which allows for more effective strategic planning since its allowance for dynamic interaction among micro and macro level factors suggests a variety of intervention alternatives. The literature of the population health school (e.g., Evans and Stoddart, 1990) has pointed out that despite exploding medical expenses in the western world in the post-war period, health indicators show declining marginal returns for such investment. The implication of this finding is that we have to rethink health policies entirely based upon medical services to focus rather on health outcomes.

The mental health system has, however, become more closely aligned with a medical care system over the last few decades. It seems reasonable to think that the limitations of the medical model for the health system in general would be even more marked with respect to questions of psychological distress. This line of thinking suggests that while medical intervention might have an indispensable role to play with respect to responding to severe psychological distress, once it has appeared, a medical model orientation is clearly inadequate for the *systems* surrounding and defining promotion of good emotional health, prevention of emotional distress, and improving the lived experiences of those in the population suffering distress.

While we may have considered such a medical orientation roughly adequate a decade ago — when many not only believed psychological distress to be an outcome of brain disease but also considered that the problem of "disease" could only be viewed from a medical perspective — it is becoming increasingly clear that while such orientation meets some "needs" it is certainly sub-optimal.

Market Model. The increased importance of this model as a narrative viewing the mental health system as "serving" patients is evident in the frequent appellation of patients today as "consumers" and of health practitioners as "service providers" or "vendors". The growing use of market concepts and language to restructure and evaluate service delivery mechanisms, including the frequent recourse to "consumer satisfaction surveys" is evident not only in the U.S. managed care context but throughout the western world.

While a desire to submit public services, including health care, to a system conceived as auto-generating efficiency and ensuring satisfaction of "consumer" tastes may seem laudable, there is in my view an excess of *blind* faith in market mechanisms as a new means of resolving problems which were formerly subject to "rational" government planning (combined, of course, with the demands placed upon governments through the political process). I do not hesitate to say "blind" faith because it is rare to see transformations of public domains toward market models guided and constrained by the assumptions contained in those models as outlined by economists dating back to Adam Smith.

The doctor-patient relation as agent-consumer? The Unbalanced paper provides a brief discussion of how a market model can be unsuited to meeting user needs, drawing especially on Goffman. Goffman's analysis of the pertinence of a "tinkering trade" model to mental health care deserves to be reviewed and updated with respect to today's mental health system, which is no longer primarily based upon asylum care.

Among the most pertinent of Goffman's comments to a mental health system "in" the community, where care consists mostly of pharmaceuticals prescribed by hospital-based psychiatrists, and occasional or chronic stays in hospital wards, is the obvious point that the practitioner is simply not the agent of the patient: the patient does not pay the practitioner, the patient has no control over hiring or firing of practitioners, and, furthermore, it is rarely the case that a patient may select among practitioners or alternative care systems.

At least as far as this thesis is concerned, the "clientele" of the mental health system is not made up of individuals such as portrayed by Woody Allen, who have the liberty of consulting at length with any manner of practitioner in order to discuss their angst, frustrations and minor neuroses. Rather, the users that this thesis and most mental health policy analysts are concerned with usually have few or no financial resources that they themselves control (they may come from comfortable families), have experienced chronic and marked psychological/emotional disturbances, and usually have "comorbidity" in terms of health and psychosocial problems and precarity.

As long as these facts characterize a mental health system, it is totally misleading to characterize the patient as a "consumer": rather, the recipient of treatment is a *patient* in the sense described by Parsons (1951) and Ingleby (1985) — the passive object of a professional practice which creates a commodity bought or ordered by persons other than the patient. If those other persons — be they doctors, family members, hospital administrators, government bureaucrats, insurance companies — are able or willing to assess user "needs" or preferences, satisfaction of patient needs accordingly does not rest upon a market model whereby consumers buy services for themselves. In other words, while "health" (or some behavioral characteristics or norms) may become commodified in a system which operates largely as a market, it is not the *market* that determines the satisfaction of patient needs but rather whatever it is that leads the true buyers in that market to buy on that basis.

Informed consent. The Unbalanced paper takes the above argument further, to the case of not only competent "voluntary" patients who in fact have little or no choice as to the commodity to be "consumed", but to incompetent patients and those for whom treatment is forced.

Firstly, it is clearly incongruent for a mental health system to simultaneously hold to discourses which view many patients as incompetent yet at the same time as "consumers". A competitive free market relies upon *informed* buyers and sellers who are aware of their own tastes and objective functions, of the variety of goods available and their prices — e.g., their costs and benefits — and who are accordingly able to maximize their satisfaction or objective functions efficiently. Such assumptions are obviously at odds with our conceptions of incompetent psychiatric patients.

Secondly, it is perverse to pretend that patients subject to forced treatment are actually having their needs satisfied as "consumers" rather than as the recipients of paternalistic assessments of what others think is good for them. While those deciding treatment may try to do so according to what they think the *patient* would have chosen if competent, as will be discussed under "best interests" below, hence *aiming* to act as the patient's agent, it is dangerous to confuse what we aim to do with what is actually happening. While the assumptions of a competitive market may allow for agents to act on behalf of buyers, their pertinence to such a market is lost if agents are not actually designated by the buyers, and also if the buyers are unable to revoke such designations[30].

Hence within current mental health systems, users, whether competent or incompetent, do not have the status of buyers, and there remains no justification for an assertion that the needs of many or most users are satisfied according to a mental health system structured as a market. I do not deny that market mechanisms could in fact advance user needs satisfaction, in constrained circumstances. Obviously, the patient would have to be competent, and have true

---

[30] The Pure Best Interests paper, to be introduced below, discusses the issue of advance directives; that paper views unquestioned reliance upon advance directives, despite the possible availability of other information indicating that the current patient's preferences might be otherwise, as in effect according a property right to the *former* patient over the later patient. Hence, insofar as advance directives are both followed and provide clear instruction for subsequent treatment decisions, we might conceive of the *former* patient as a "consumer", but not the *present* patient.

consumer sovereignty. Consumer sovereignty implies a number of assumptions regarding a competitive market, including a variety of alternatives, freely available information, and competitive rather than oligopolistic or monopolistic supplier structures. None of these assumptions apply fully or even partially in mental health care today.

The above discussion, drawing mostly upon the Unbalanced article, suffices in my view to reject the market model as a credible narrative that user satisfactions are met today, while also suggesting the theoretical applicability of the narrative under a number of important constraints. There are two other points which I will raise here which cast doubt upon whether a market model can *ever* be appropriate for the functioning of a mental health system. The first relates to health as a public good. The second raises welfare economics questions important for a meta-evaluation of a market economy as a means for social allocation of goods.

Health as a public good. The collective efficiency of the market, and hence the degree to which it is an appropriate mechanism for satisfying user needs according to a market model, requires that the goods bought and sold are *private goods* in the microeconomic sense: that there are no consumption externalities in that costs and benefits of consuming goods are borne *only* by the buyer. The way in which "health" is conceived in the Pop-Health paper implicitly suggests that this is far from the case: health and ill-health have many social determinants that can never be excluded from etiological processes, and health and ill-health outcomes have wide-ranging impacts on other individuals as well as upon the functioning of societal systems including the economy.

What this suggests is that if health, or emotional health, is a public good, individual purchases of "health goods" may be too few or too many, thereby failing to optimize the allocation of health goods with respect to other goods at the *collective* level[31].

---

[31] Whether purchases of health goods — i.e., services or products aimed at preserving or improving health — will be excessive or insufficient depends upon the respective allocation of externalities with respect to the costs and benefits of those goods. Hence liver transplants may be overbought when the recipients of the livers do not have to pay their full costs — e.g., if they are taken without consent or compensation from infant children. Sports club memberships may be

Welfare economics. The second point raises an issue applicable to the evaluation, from a collective perspective, of *any* market as a mechanism for achieving social welfare. In economics, the subfield of welfare economics, which considers the impact of *distributions of wealth* upon the functioning of markets, has been the poor cousin among specialties, and has been almost completely ignored in the current neoliberal era — an era whose theories of economics are, in my opinion, exceedingly simplistic and one-sided (biased in favour of *laissez-faire*, regardless of the economic models invoked).

Attention to welfare economics questions, when evaluating the outcomes of a particular economic model over time, might reveal, for example, that few people are actually able to participate significantly in a market-place otherwise considered "efficient", due to lack of capital. Unless this is realized, we may fail to appreciate the inappropriateness of the current economic structure or its potential to be adjusted to achieve better collective outcomes[32].

Furthermore, an economic structure may lead to increasing disparities in the distribution of capital over time, leading to fewer and fewer individuals able to achieve satisfaction through the market place. Unfortunately, theoretical evaluations of outcomes in economics tend to be based on static models which do not show dynamic change in these models over time whereby the nature and number of players, and their "endowments" (what they bring to the game) can be changed.

More seriously, few economists today are willing to work with criteria that allow for meaningful evaluation of economic systems from a *collective* perspective. As discussed in the part concerning methodology, given the impossibility of objectively quantifying "utility" (happiness, satisfaction), of making interpersonal utility comparisons and hence aggregating utilities to provide collective measures,

---

underbought in a society where the more fit people are, the more they make *others* happy and productive.

32. One might think such a blind spot too large to be missed, yet it seems evident that the obvious is very often rarely stated, for a variety of reasons. For example, I agree with Richard Wilson (1997) that poverty and unemployment is more frequently explained today in terms of the failings of individuals than in terms of the basic economic structure of our society.

economists rest with evaluation criteria which can be *measured,* but are of questionable *value* and *validity* with respect to human welfare[33].

It is not necessary for the purposes of the present discussion to go into my view that utility theory is indispensable for appropriate evaluation at the collective level of economic policy choices — suffice it to say that while a market model might be *efficient* in given circumstances, in that most people participating in the market get good deals that at least in a short term analysis improve their lot, it might be sub-optimal at the collective level, in that many or most become worse off or improve far less than some minority.

Poverty and mental health. This discussion of economic theory is more pertinent to the mental health system than might seem apparent. The social welfare problems of markets decrease insofar as the traders in those markets start off with roughly equally valuable goods to offer for trade, and insofar as the market does not contain a dynamic that increases endowment inequality over time. Yet the vast majority of mental health users are poor; while some such poverty may be due to psychological distress, it is also obvious that poverty itself augments psychological distress and the risks of its onset.

Indeed, insofar as "the" mental health system meets user needs *because* of market operations, it is with respect to individuals who are fairly wealthy, and are able to purchase from a selection of alternatives including psychotherapy, various types of counseling, home nursing and accompaniment with daily living functions and, in private health systems, the best hospitals, rest homes, and medical specialists. But unless governments are prepared to address problems of income distribution or poverty *a priori,* or at least establish an extensive voucher system (which is probably a less satisfactory solution in that it may lead to overconsumption of vouchered goods — e.g., medical — over non-vouchered goods — e.g. housing — when the latter may also be important or crucial with

---

[33] E.g., gross national product, which ignores who has what and incorporates economic activity providing little or even negative collective welfare (machine gun sales to biker gangs are included, destruction of electrical distribution infrastructures are not subtracted but replacement of destroyed infrastructures is), and Pareto optimality, an efficiency criterion which is reached when no person could get a better deal, given what they have to trade with.

respect to user needs and health outcomes), it would be perverse to consider a market system as a reasonable way to satisfy the needs of those suffering psychological distress, when such markets require buyers with money yet poverty is widespread and systematic among distressed persons.

The pertinence to user needs of a market model. To conclude this discussion of the market model as a narrative providing an expectation that user needs are met, the above arguments suggest that not only would such a narrative be misplaced in our present mental health systems, but that due to the assumptions of a competitive market, such a market could only be made applicable to severely constrained circumstances and clients.

Furthermore, if health is a public good, markets should never replace completely government or collective intervention to correct distortions due to externalities, and to invest in health to the extent that markets will systematically under-invest in health. Finally, given that poverty and mental health are intertwined — we do not yet understand how, since there has been far less research on this relationship than there has on the etiological role of biological factors — a crucial prerequisite to a market solution must be addressing such poverty, or otherwise providing emotionally distressed persons with the resources that can enable them to make meaningful choices.

Best Interests Model. The Unbalanced paper distinguishes the market service provider from the person making decisions on behalf of the patient under the best interests model, in that the best interests decision maker possesses authority to select *outcomes* for the patient, whereas the delegated authority of the market service provider is restricted to the technical issues of creating the outcome selected by the client. This is because the best interests model implies selecting not what the patient *wants* but rather what the patient *needs*, or, more generally and also rather more ambiguously, what is in the "best interests" of the patient.

Influence of the best interests model as narrative for meeting needs for all patients. I think it is reasonable to suggest that the narrative most influential in leading people to believe that the mental health system can or does generally meet user "needs" is implicitly provided by the best interests model, even though the bioethics literature restricts the applicability of this model to substituted decision making on behalf of incompetent patients, psychiatric or otherwise. The influence of this narrative with respect to psychiatric patients is more obvious when applied to those designated incompetent, since it is invoked in contexts of imposed treatment for those who "need" it but are viewed as too irrational to realize that it is for their own good.

Despite the ascendancy in the ethics literature and jurisprudence of "substituted judgment", when there is good evidence of what the patient would want if competent, or advance directives when there is proof of what the patient, when previously competent, instructed to take place in the event of subsequent incompetence, there seems little doubt that in cases of psychiatric care, where it is rare to have patients who are unable to communicate preferences (whether or not they are "true" or rational preferences), forced treatment is justified by appeal to a paternalistic assessment of what is beneficial for the patient.

As a theoretical concept, "best interests" remains very vaguely as well as variously defined, allowing for various nuances which may or may not be concretely differentiable, e.g., what the patient needs, what the patient "really" wants, what is good for the patient, etc., all of which provide a seeming rationale for forcing a treatment upon patients who say no but whose refusal is considered irrational and/or the result of incompetence. I suggest that such conceptions largely underlay an implicit confidence in the operation of the mental health system with respect to psychiatric patients in general, not only those formally designated as incompetent, for a number of reasons.

Mental malfunction, irrationality, and incompetence. Many people probably view psychiatric patients as at least somewhat irrational, by *definition*, even if only the least rational or those refusing treatment will be formally found to be incompetent. Despite the lack of good evidence that the relationship between psychological distress and irrationality is more than tenuous, and the existence of alternative theories that both "mental illness" and "irrationality" are, when it comes to the vast majority of psychiatric patients, labels applied to certain forms of socially disapproved deviant behaviour, the two concepts are commonly confounded.

The metaphor of a "malfunctioning brain" is taken literally by many people who view the brain as some kind of machine which orders meaning and expression, and which has had a mechanical breakdown (or, not dissimilarly, a "biochemical deficit"). It is logical for those who believe that persons who have been given psychiatric diagnoses are suffering from such a malfunctioning brain to then conclude that the person is necessarily at least somewhat impaired in terms of rationality and hence competency. This helps to explain why formal competency assessments usually ensue only when treatment is refused: treatment is viewed as repairs to a broken brain which would only refuse such treatment because it is broken and does not know any better.

As a narrative underlying faith in the mental health system, therefore, a best interests justification may seem inextricably intertwined with a pervasive medical model ideology. However, the two models can be separated as justificatory narratives insofar as concepts of patient incompetence have some logical and empirical independence from concepts of irrationality tautologically related to a diagnosed mental illness. If indeed we are looking at the latter situation, the justificatory narrative reverts to the medical model, which as discussed above does not provide a credible justification, unless we are prepared to view medical practitioners and what they do in terms of a metaphysically based meritocratic theory of society.

If we view some level of patient incompetence, at least partly independent from psychological distress, as a prerequisite for the appropriate application of the best interests model, then of course we are obliged to recognized that *a best interests narrative, as a hypothesis that user needs are met, can only pertain to patients with some level of incompetence, whether formally assessed or not.* The rest of this section will therefore discuss the acceptability of the best interests model as a narrative supporting an assertion that the needs of patients with apparent impaired competence (hereafter referred to as incompetent) will be met.

Evaluation of best interests model on behalf of incompetent psychiatric patients. As indicated in the Unbalanced paper, justification for proxy decision making under this model must be predicated by an accurate, unbiased objective assessment of incompetence, and, if that condition is met, be supportable on ethical grounds in view of the primacy western societies accord to self-determination and the dignity of the person.

The validity and reliability of incompetence assessment is a problem well-recognized in the literature; one of the severest problems being the independence of assessment of incompetence from a diagnosis of mental illness and the refusal of treatment (as is discussed in papers to be introduced below).

The ethical value of beneficence on behalf of frankly incompetent persons, expressed in jurisprudence over the State's role as *parens patriae*, has been commonly asserted as a value either not inconsistent with values related to the liberty of the person, or which must inevitably be weighed against such values in order to achieve "good" outcomes on behalf of competency impaired persons. This thesis takes issue with this view, as will also be discussed in papers to be introduced below.

However, this thesis has for the most part avoided the competence assessment condition, and will come to the issues of ethical justification for proxy decision making on beneficence grounds after having analyzed the best interests model within its own logic. Even assuming that a substituted decision making process has satisfied competency assessment conditions, and leaving aside the beneficence ethical justification for proxy decision making, there are two further necessary conditions which must be satisfied before a best interests model can achieve credibility as a narrative asserting that the needs of incompetent patients are met.

The mere existence of incompetence does not suffice to invoke decisions on behalf of an incompetent patient under a best interests model based upon beneficence. A proxy decision making process under such model has to have some reasonable prospect of actually attaining its beneficence objective, i.e., improving the patient's well-being. This implies the two necessary conditions, described in the Unbalanced paper, that *the proxy process must be able to actually determine what would serve the patient's best interests*, and that *it must be designed such that proxies will actually try to determine those interests and satisfy them.*

Validity Problems. Evaluation of a best interests model is impossible if we do not know what in fact the objective of the model is. The article "Toward a Pure Best Interests Model of Proxy Decision Making for Incompetent Psychiatric Patients" (the "Pure Best Interests paper"), written with David Weisstub, argues that there has been no clear definition of "best interests", that interpretations of a best interests criterion have varied widely, and that in both practice and theory many of the interpretations conflict with a beneficence criterion, or operationalization of this criterion in the best interests model has conflicted unnecessarily with patient liberty interests.

That paper therefore proposes a revised model for proxy decision making for incompetent psychiatric patients, the "pure best interests model". This revised model contains a clear and operationalizable criterion: to make the decision that the patient would have made if temporarily competent.

Please refer now to the article "Toward a Pure Best Interests Model of Proxy Decision Making for Incompetent Psychiatric Patients".

Under the best interests criterion, there are few possible grounds to support a narrative that patient needs are met, because there is no clear theoretical or empirical relationship between how proxy decision makers are supposed to determine "best interests" and actual patient needs. As indicated in the Pure Best Interests paper, the best interests criterion has been allowed to incorporate not only paternalistic or biased assessments by care-givers or proxies regarding what they think is "good for the patient" but has also been allowed, implicitly or explicitly, to incorporate non-patient interests — e.g., those of the State, family or care-giving institution.

Furthermore, there is no good reason to found a proxy decision model in the ethical principle of "beneficence" to the detriment of the patient's own liberty interests. If "incompetence" truly reflects *incapacity* to function as a decision maker adequately satisfying one's own interests and needs, proxy decision making should seek its ethical justification in an aim to improve the patient's capacities to pursue those interests and needs.

It should therefore be seen as an *empowering* process which, insofar as proxy decision making is necessary, supplements the patient's existing capacities and aims to further improve them. A proxy decision process which further infantilizes patients and reduces their abilities to pursue their own interests as autonomous persons would therefore not be justifiable under the Pure Best Interests criterion, which is based *solely* on autonomy values.

A proxy process based upon *beneficence* is not based solely upon the patient's own needs, whatever they are. It might be reasonably argued that a pure best interests proxy process based upon autonomy values represents an operationalization of incompetent patient needs assessment and satisfaction, insofar as it is believed that what a competent person decides is in accordance with (or at least strongly related to) that person's own needs, and that a proxy process which at least aims to replicate what the incompetent patient would decide if temporarily competent might hope to validly (albeit with reliability and bias problems, to be discussed below) estimate what competent patients would decide.

However, the objective in developing the pure best interests model was not to better meet patient "needs" *per se*, but to firstly establish that best interests models (as well as substituted judgment and advance directive rules or criteria) have a problematic relationship to patient needs however conceived (e.g., interests, expressed preferences, basic values, long-standing preferences or tastes), and secondly to establish a theoretical objective for substituted decision making which can allow for evaluation of proxy processes based solely on the autonomy values of the patient.

Again I emphasize that I have never assumed in this thesis that the very concept of patient "needs", *per se*, should have a place in the development and evaluation of a mental health system. Nevertheless, the pervasiveness of existing explicit and implicit narratives justifying the system as centered around patient needs being met still need to be analyzed in order to clarify whether such statements have any meaning, if so, what meaning, and finally whether needs or some other concept, perhaps related, should be more pertinent for evaluation of a system insofar as it is supposed to be "for" users.

Proxy decision making under uncertainty. Regardless of the rule or criterion invoked in proxy decision making, the process and/or results are potentially evaluable according to the pure best interests criterion, in terms of its likely or actual success in making the decision that the patient would have made if temporarily competent. This could be done even for a proxy process that does not pretend to direct its decisions to the interests, needs or underlying desires of patients (e.g., a "dummy" system, which refuses to make any decisions, or a system which provides unquestioned authority to family members to decide whatever they want). It is not necessarily the case that a particular proxy process with a good aim must be better than another process with a bad aim, in terms of its good and bad outcomes over the long-run.

The manuscript "Error under a Pure Best Interests Model of Proxy Decision Making: Implications for the Justifiability of Forced Treatment" (the "Error paper") was originally written in order to make evaluation of proxy decision making possible, by outlining evaluation criteria regarding reliability and bias. The Pure Best Interests paper grew out of the Error paper, once it became apparent that evaluation of proxy decision making was impossible when it was not clear *what* proxy decision making was supposed to accomplish (other than coming up with a decision). Once a pure best interests criterion is applied, problems in actually satisfying that criterion can be addressed, and alternative proxy processes might be evaluated according to issues of validity, reliability and bias.

Please refer now to the manuscript "Error under a Pure Best Interests Model of Proxy Decision Making: Implications for the Justifiability of Forced Treatment".

It should be stressed that the purpose of the pure best interests "model" is not to define a particular proxy *process*, but rather to enable evaluation of any proposed process with respect to how successful it is in making the decision that the patient would make if temporarily competent. It may be that it would be difficult to establish that *any* process would meet that criterion — but to the extent that the given criterion is the only justifiable basis for proxy decision making, or for replacing a patient's expressed wishes with the views of another, the appropriate response to such difficulty is not to adopt a process which easily meets another criterion whose ethical justifiability is questionable, but rather to honestly ask ourselves why we are making decisions for other persons (particularly when those other persons are able to express a preference).

The Error paper outlines the many ways in which uncertainty enters into proxy decision making — an issue which has rarely been discussed in the literature with respect to substituted decision making in general, and almost never with respect to psychiatric patients. Yet the implications of uncertainty are important for the ethical justifiability of substituted decision making on behalf of psychiatric patients, because virtually all of such patients are able to express a preference — whether or not they are "correct", rational or competent.

Among the problems increasing uncertainty as to whether proxy decisions or decision processes succeed in making the decisions that the patient would have made if temporarily competent, as discussed in the Error paper, are:

- Bias: Self-Confirming Hypotheses. This is one of the most fundamental problems in decision making on behalf of psychiatric patients designated incompetent, in that very often competence assessments and proxy processes are invoked *because* the patient declines a recommended treatment — which is then confirmed in the decision made on behalf of the patient.

  Insofar as proxy decisions are tautologically related to treatment refusal, the bias problem becomes magnified to one of validity: the proxy process is not trying to achieve the decision that the patient would have made if temporarily competent, but rather the decision favoured by caregivers.
- Bias: Sociodemographic Differences between Proxies and Patients. This is not an occasional issue, whereby different biases tend to be minimal or cancel themselves out over the long-run, because those involved in decision making on behalf of psychiatric patients designated incompetent tend to be markedly different from such patients in terms of education, social class, income and wealth, race, and gender. These distinctions make it difficult for proxies to put themselves in the patient's shoes — if in fact they really try to do so rather than substitute their own values.
- Bias: Conflicts of Interest. There seems little doubt that there are serious conflicts of interest between proxies and patients; see, for example, the discussion of interest divergence in the Unbalanced paper. Proxy processes could be set up to minimize the impacts of such conflicts, or to provide decision making authority to disinterested persons, but this is rarely the case. In most jurisdictions, the treating psychiatrist's recommendations are unquestionably followed where those with nominal proxy authority are not the treating psychiatrists: i.e., family members, treatment review boards dominated by medical practitioners and institutional interests. Treatment refusals by patients considered incompetent are so rarely upheld that it is quite reasonable to ask whether the patient, even if frankly incompetent, might nevertheless be expressing reasonable fears and basic values in refusing a particular treatment — which

treatment is often refused because of iatrogenic effects and other impacts on the patient's autonomy and quality of life.

- Reliability and/or Bias: Attitudes to Risk. Microeconomic decision theory has long recognized that attitudes to risk — based on consumption utility functions comparing alternative distributions of goods over time — vary among individuals, and can easily lead to very different choices even among people who have the same tastes with respect to outcomes and the same assessments of risk (i.e., assessments of the probabilities of various good and bad outcomes given various decisions). It is almost never acknowledged in the substituted decision making literature on behalf of psychiatric patients that the values of the patient are pertinent in this respect.

  This means that even if risks can be objectively determined by a proxy decision maker, and the costs and benefits to the patient of various outcomes can be accurately estimated, replication of the decision the patient would have made if temporarily competent also requires taking into consideration issues like whether the patient is prepared to pay a cost now (e.g., a treatment with temporary but very onerous side effects) in order to be probably much better later on.

  Insofar as proxy decisions that fail to account for the patient's own risk utility function have "random" error distributions centered upon the aggregate patients' risk utility functions, the series of proxy decisions will have questionable *reliability*.

  Insofar as proxy decisions incorporate, in fact, *proxies'* views of what would be appropriate risk utility functions for patients, or indeed proxies' *own* risk utility functions, a systematic *bias* could arise in proxy decision making. For example, it would be interesting to test the hypothesis that proxies are more willing to gamble — pay a price now for a possible larger benefit later on — than are patients. An *a priori* argument for such hypothesis (admittedly not sufficient) might be that it is obvious that in cases of proxy decision making on behalf of *psychiatric* patients, temporarily disabling treatments (e.g., neuroleptics, electroshock) are much more favoured by proxies than by those incompetent patients expressing an opinion.
- Bias: Controlling Risks. A dynamic perspective of decision processes recognizes that part of a rational decision maker's calculus is the degree to which decisions

foreclose future options, and the degree to which, accordingly, the decision maker could make future adjustments, change directions and, in effect, *alter* risk probabilities by intentional action over time. Hence decision makers prefer risks that they can control, providing them with the flexibility to continuously re-evaluate their circumstances in comparison with their objective functions (which may also change over time).

I have never seen this issue directly raised in the psychiatric literature, although it is implied in Winick's many discussions of the therapeutic benefits of autonomy. I would suggest that to the extent that the importance of this factor is ignored in substituted decision making with respect to psychiatric patients, an important bias is created, for the simple reason that the kinds of decisions usually favoured by proxies tend to *reduce* the patient's flexibility over time to re-evaluate circumstances and possibly change course. Neuroleptic drugs commonly produce passivity and reduce patient initiative, which is one reason often given for refusal of such treatments (yet, some critics have argued that such outcomes are the *objective* of the treatment). Electroshock induces short-term memory loss, and sometimes long-term as well. "Voluntarily" signing in to a hospital or asylum for treatment often means an obligation to stay for a fixed period, or until the treating psychiatrist decides to let the patient go.

- Bias: The Impact of Coercion. There is increasing recognition in the ethics literature concerned with competence-reduced persons (elderly, persons with intellectual deficits, persons suffering severe psychological distress), and a gradual acknowledgment among community care services, that those with decision making authority should favour as much as reasonably possible the *expressed* preferences of incompetent persons, even if those decisions are clearly "wrong", if the apparent negative impacts of the decision upon the person's welfare are not too important.

This recognition reflects a growing awareness of the negative emotional impact of coercion upon vulnerable persons; in other words, the benefits of a decision that would be favourable for the vulnerable person if that person had made the decision can become negative once the person sees the decision as forced. Under the pure best interests criterion, the proxy should try to imagine how the patient would decide when *temporarily* competent — i.e., taking into account, as the temporarily competent patient would do, the subjective impacts of the decision once that person lapses back into incompetence. I would suggest that unless proxies make a conscientious effort to reduce coercive decisions to a minimum, their decisions on behalf of incompetent persons will be biased.

All of the above provide means for evaluating proposed mechanisms for proxy decision making — helping therefore to develop improved processes — but also suggest that proxy decisions could be wrong far more often than we might have realized under implicit assumptions that one person can make good decisions on behalf of another.

There are a number of possible implications of proxy decision making uncertainty for how decisions might be made, some of which are discussed in the Error paper in the section concerning decision rules. There is perhaps no one precise way to handle all types of patient or circumstances, but, in accordance with decision theory, the method used should in some way take account of the consequences of error.

For example, is it better, from the patient's standpoint, to wrongly unplug a respirator or wrongly leave it plugged in? In end of life issues a consensus has formed that when in doubt the patient should be kept alive, unless the patient is suffering severe pain. Such a rule of thumb might be justified on both ethical and decision theory grounds. Another perspective to guiding the selection of decision rules, inspired by Rawls' theory of justice (1971), is to try to use rules that the *patient* would want used.

In the case of psychiatric patients designated incompetent, if the value of autonomy is not only paramount but the only guiding value for substituted decision making, an appropriate rule of thumb would be to honour expressed preferences unless an alternative decision can be justified as consistent with the patient's autonomy values. It might be argued that substitution of an alternative for expressed preferences should meet a standard of "reasonable certainty" that the decision the patient would have made if temporarily competent has actually been estimated. This decision rule would provide weight in the decision for "autonomy", in order to make a decision one way or another in the face of uncertainty.

In practical reality, however, this may not provide a high enough threshold, insofar as the proxy decision making process has still incorporated a number of biases as outlined above, which all tend toward a decision which would counteract expressed preferences[34]. If this is indeed the case, as much of the evidence drawn upon in this thesis seems to indicate, I suggest that the burden of proof that the needs of incompetent psychiatric patients are met via proxy decisions rests with those making such an assertion. There is little evidence in the literature to support such a position.

Estimates in the literature of proxy decision accuracy. There are very few estimates in the literature of the accuracy of proxy decision making. The few empirical estimates which exist concern non-psychiatric patients, and may use various indicators of what the patient would have decided that might be considered more or less reliable and valid for the purposes of evaluation studies pertinent to a pure best interests criterion (methodologies include advance directives not revealed to dummy proxies, experiments involving persons actually competent, analysis of reasons and criteria given by proxies).

A quick overview of that literature, as provided in the Error paper, suggests accuracy which is only fair to poor. Given that far more attention has been paid to proxy decision making on behalf of geriatric patients, minors, end of life issues regarding unconscious persons, and persons suffering intellectual deficiencies, than to psychiatric patients, and that the proxy decisions on behalf of those persons tend to be viewed as much less "medical" issues than are proxy decisions on behalf of psychiatric patients, which usually concern administration of psychoactive drugs or electroshock, I think it is prudent to assume that existing accuracy of substituted decision making on behalf of psychiatric patients is even worse than "fair to poor": i.e., failing grades.

---

34. My view is that public policy should prohibit *psychiatric* treatment against the expressed preferences of patients, on the grounds that no *psychiatric* treatment has a benefit-cost ratio so strong, on average, that a valid, unbiased assessment of attributed patient preferences could ever, with reasonable certainty, overcome expressed patient preference. This view is not defended in the thesis, however, since the thesis deals only superficially with the benefits and costs of psychiatric treatments. Note that this position does not exclude *any* forced treatment, only forced *psychiatric* treatment. I do not exclude, for example, the possibility of forcing life-saving surgery upon clearly incompetent persons whose prospect for quality in life is good, given a variety of conditions.

Again, whether or not one would be comfortable with such an assessment, there is really nothing in the way of scientific evidence to suggest that the needs of incompetent psychiatric patients are met, insofar as their "needs" are strongly related to what those patients would have decided if temporarily competent. There is much that could be done, on the other hand, that might advance the satisfaction of the needs of these patients; we need to start with clear evaluation criteria with justifiable ethical rationales, to assess existing processes of proxy decision making and improve them accordingly. The "best interests" narrative does not provide such criteria, leaving it a weak narrative with respect to incompetent psychiatric patients (and, as noted above, inapplicable with respect to competent patients).

The Power Access Condition *(The Political Model)*

Even if non-user actors pursue interests whose attainment would infringe on the satisfaction of user needs, and even if policy-makers, government bureaucrats, and service administrators would be inclined to be disinterested in reforms and other interventions upon the mental health system which would advance user needs, it might still be possible to construct a mental health system that satisfies user needs insofar as users themselves hold sufficient power, at individual, mental health system and/or political levels, to achieve such a result. This latter possibility is derived from a political model of society which asserts that individuals seek to satisfy their interests through the exercise of power.

Hence the "power access condition" provides a candidate narrative for the hypothesis that user needs will be met by or within the mental health system to the extent that they have power compared to other players. Note that this condition must, in order to satisfy that hypothesis, imply that what users seek through the use of power is what they "need". Also, in its simple form, this condition implies that the "game" is zero-sum, in that successful exercise of power by one group over another increases the "goods" for the first group to the detriment of the second group.

Interaction with the Interest Convergence Condition. Of course, in the real world, not all exercises of power are zero-sum, *insofar* as there is convergence of interests among the actors which, as discussed above, implies either that satisfaction of one player satisfies the other (either because the goods are public goods collectively consumed, or because the game is positive sum: pursuance of separate interests increases the supply of goods, to the benefit of all).

For analytical purposes, however, the issue of interest convergence can be separated from that of power access. Hence, for example, we might think that a "market" would provide interest convergence, resulting in a positive sum game whereby those with deficient power might still have a good share of available goods — growing because of market competition. Depending on the nature and degree of the convergence and its implications for a game's dynamics, more or less power access might be required in order to obtain x amount or x% of available goods. A market narrative might necessitate some minimal amount of power (capital, information, mobility, choice, autonomy, capacity to communicate to potential buyers and sellers) to achieve some adequate level of need satisfaction, whereas a best interests narrative might require no power at all, insofar as those who do have power operate to charitably "give" to the powerless.

Hence meeting each of the interest convergence and power access conditions is not a matter of all or nothing, as some limited degree of convergence combined with some limited degree of power access might be sufficient, depending on the models referred to, for obtaining a reasonable or ethically justifiable share of society's goods.

Evaluating the power access condition was by far the easiest task of this thesis; all that was required was to state the obvious. On the rare occasions when user power is discussed in the therapeutic or policy literature, it is acknowledged that the typical psychiatric patient has little power in the therapeutic situation and virtually no power with respect to mental health policy and administration.

Unfortunately, there is a dearth of political science analysis of the politics of mental health (with the notable exception of the work of David Rochefort), but sociological (e.g., Mechanic, White) and anthropological (e.g., McLean) descriptions of mental health systems show (often implicitly, in that users are not included among discussed stakeholders or actors) that "participation", "democracy", "partnership" and "community activism" in mental health systems exclude users, whose role is greatly overshadowed by a variety of "stakeholders" including family members, professional and volunteer caregivers, "community resources" (usually meaning more or less "professional" caregivers located "in the community" as opposed to in hospitals and asylums), health and hospital administrators and practitioners.

A Case Study: The Right to Treatment. The Unbalanced paper, in the "power disability" section, noted that most literature which discussed the role of users in attempting to affect change in the mental health system has focused upon their efforts to achieve improved legal rights, but that it is difficult to separate out the impact of user groups with those of other actors who had converging interests with users.

The manuscript "'Meeting the Needs of the Mentally Ill': A Case Study of the 'Right to Treatment' as Legal Rights Discourse in the U.S.A." (the "Right to Treatment paper"), written with David Weisstub, provides a case study of this issue, by analyzing how the first of the major "rights" claims, with respect to psychiatric patients, managed to receive some attention. This paper is pertinent to this part of the thesis because it shows how little of rights advocacy was actually in the hands of users.

Please refer now to the manuscript "'Meeting the Needs of the Mentally Ill': A Case Study of the 'Right to Treatment' as Legal Rights Discourse in the U.S.A."

An historical perspective of the "right to treatment" shows that this was actually a game of lawyers, judges, administrators and psychiatrists from insane asylums. After many years of litigation — lobbying at the political and legislature levels appears to have been inefficacious (although exhaustive review of the literature reveals little which discusses this point of attack) — the limited support of civil rights lawyers for the few users who were able to obtain access to the courts was not sufficient to accomplish much more than living standards in institutions equal to those of criminals along with a treatment regimen left to the "professional judgment" of institutions and practitioners. The litigation failed to achieve constitutional recognition that patients had the right to the "least restrictive" treatment, and consequently the services that would facilitate independent or community living and allow institutionalized patients to regain some measure of autonomy.

Indeed, not only did the "right to treatment" become an almost empty right to minimally humane standards for maintenance in institutions, the discourse of patients and their advocates was coopted by proponents of *forced* treatment who asserted that many psychiatric patients who refused treatment were too irrational to realize that they *need* it, but since they had a *right* to it society was obliged, on beneficence grounds, to impose it.

Hence the right to treatment issue encapsulates well the findings of this thesis to the effect that in the contemporary mental health system, patients themselves have little power, and that if their needs are to be met it must be on beneficence grounds — which grounds are basically impossible to evaluate.

Rights and empty rights. The Right to Treatment paper concludes by pointing out that "rights" in our societies are empty unless the holders of those rights have the power to pursue and enforce them. Furthermore, the *parens patriae* obligations of the State on behalf of vulnerable persons should be understood as based on autonomy rather than paternalistic beneficence values; i.e., that the objective of State help and care is to facilitate the development of the vulnerable person's autonomy.

Such a criterion is far clearer — and more operationalizable for policy and evaluation purposes — than an obligation which merely requires doing what is "good" for the person. A "right to treatment" interpreted under such an enlightened conception of *parens patriae* would then imply assisting the person with what is required to progress toward independent or normalized community living, which of course means making one's own decisions as much as possible and pursuing one's own interests — hence, *empowerment.*

Empowerment. Empowerment, as discussed in the Pop-Health paper, is a concept which both recognizes the therapeutic benefit of achieving a sense of personal and social mastery and the necessity of providing real social power in order to obtain this sense. Empowerment theorists view the development of psychological distress as at least partly caused by feelings of powerlessness, often related to a marginalized position in society. The therapeutic jurisprudence literature suggests that empowerment is important for ameliorating psychological distress — a position which is quite consistent with a major etiological role for somatic factors.

Sources of systematic disempowerment. Yet psychiatric patients are *systematically* disempowered. The Unbalanced paper outlines a variety of sources of power deficits relative to others in the mental health system, in their families and in their communities:

- the passivity engendered by the "*patient role*", which is expected by society and encouraged by current therapeutic practices;
- poor self-confidence and/or a reluctance to engage in collective action due to the enduring cultural *stigma* of mental illness, which impacts upon how users are perceived by others and how others listen to them, which imparts feelings of shame, humiliation and inferiority to users who may share the same stigma, and which militates against activities which might expose the identities of users as recipients of psychiatric care;
- the direct effects of psychological *distress*, whose impacts upon the capacity of individuals to pursue their own interests in competition with others may be less due to irrationality or incompetence because of a malfunctioning brain than to the confusion, lassitude, fatalism, lack of self-confidence, difficulty in concentrating, emotional and perceptual distortions and other conditions that can be associated with highly stressful situations (e.g., ordinary medical illness, extreme poverty, emotional or physical abuse);
- the *iatrogenic* effects of somatic and non-somatic treatments, which on a psychological level can be infantilizing and on a somatic level disturb normal brain functioning, affecting emotion, cognition, and movement — indeed, it is an open question in the literature whether neuroleptic treatments are more disabling for many persons diagnosed as schizophrenic than the condition being treated;
- structural disadvantages compared to other groups of actors who because of money or State-derived authority have been able to overcome *free-rider problems* with respect to collective action — e.g., professional organizations and unions which can enforce membership, and pharmaceutical companies in an industry which has no free-rider problem because it is oligopolistic;

- structural disadvantages in access to the *monetary resources* which enable or facilitate collective and individual action in pursuit of user interests, in that while most users are poor and user organizations receive little or no government funding, profits enable corporations to pursue their interests, psychiatric associations receive membership dues from highly paid practitioners as well as promotional revenue from drug companies, and family lobby groups receive not only membership contributions from members with greater wealth and earnings than users but also various subsidies from governments and pharmaceutical companies. Some of these funding sources not only provide major advantages to non-patient actors but can also be expected to alter their interests to align more closely with those of medical and pharmaceutical interests — thereby affecting the degree to which interests converge among users and non-users.

In sum, user power to influence the mental health system is not evident in the structures of the system, which provide little or no role for user participation in determining policy, nor is user power evident in historical studies of major transformations and reform initiatives, and, finally, nor can we *expect* users to have an important role, whether at the policy, service delivery or therapeutic levels, given the systematic nature of their power deficits compared to other actors. Not only is there little evidence to support a hypothesis that user needs might be met according to the power access condition, it seems clear that there is more than enough evidence to firmly *reject* such a hypothesis.

Indeed, I think it fair to say that insofar as any of the interest convergence conditions require any significant level of user power — even fairly small — to provide sufficient support for the overall hypothesis that user needs are met, those conditions as well should probably be regarded as inadequately met. This is the case, for example, under the market model, which even insofar as there truly seems to be a "market", users have insufficient power to act as the "informed consumers" that a market model requires for efficiency and meeting the needs of consumers.

Conclusion:
A Critical Political Economy of the Mental Health System

Discussion of Results

This thesis addresses research questions deductively generated by the overall hypothesis that there are reasonable grounds to believe that the mental system will meet the "needs" of users. Note that attention is paid to the micro (patient / substituted decision makers), macro (policy / politics over time), and meso (partial analysis of sub-systems and their interactions by looking at patients and psychiatrists) levels of a broadly defined historical mental health system.

When this thesis was conceived I had not expected it to prove or disprove the theory constructed for the thesis, or its counter-propositions; the primary objective was to provide a framework — with enough flesh to show its usefulness — for analysis and evaluation of the mental health system, and hopefully provide new directions for planning reform of the system. Nevertheless, the analysis has ended up firmly rejecting some candidate narratives, and those remaining are so constrained or questionable that I will conclude that there are no reasonable grounds to accept a hypothesis that the mental health system meets user needs.

Below I will summarize what the thesis has found with respect to the interest convergence and power access conditions, explore their implications with respect to the overall hypothesis, and then discuss the applicability of the policy intervention condition, given the interacting limitations of the other two.

Interest Convergence Condition. The thesis set up three narratives which might be appealed to in support of an assertion that the mental health system meets user needs because the interests of those with power converge with those of users: the medical, market, and best interests models.

Medical model. The medical model was found not to be sufficient as such a narrative, even if we were to accept the role of somatic factors in psychological distress. User needs would normally be conceived to go well beyond the medical, and even "medical decisions" must incorporate a wide variety of factors — user valuations of costs and benefits of various outcomes — that go far beyond purely medical expertise.

Insofar as somatic factors are not the most important in etiology of psychological distress, and of lesser importance than psychosocial factors for alleviation of distress, a medical model becomes even more misplaced as a useful basis upon which to build a mental health system aimed at meeting user needs.

Furthermore, at system-wide levels we need to recognize that even if somatic factors seem key and immediate for individuals developing diagnosed mental illnesses, prevention of such illnesses at the *population* level requires focusing upon a variety of system-wide cultural, economic and social factors which facilitate the development of psychological distress, interact with somatic factors, and impede amelioration.

Given these considerations, while medical treatment as a key element in the mental health system might be justifiable on various grounds including efficacy and, *accordingly*, user choice, confidence in the system *because* it is medical would seem to be based more upon a meritocratic view of society which views medicine and medical practice as good by definition, rather than upon some conception of the needs of those persons suffering psychological distress.

Indeed, this is the view of some of the critical literature which, remarking upon the lack of a scientific basis for biopsychiatric claims, explains society's confidence in a biopsychiatric approach to the mental health system in terms of, in effect, a medical meritocracy or psychiatry as religion (control of deviance might be seen as the flip-side of this perspective, in that those with status need to "construct" and control deviance to justify their social role).

Market model. This narrative comes nowhere near to providing a convincing argument that user needs might be met. Firstly, it is rare that mental health services are structured as a competitive free market; patients are given few choices and "service providers" are not the agents of patients, who do not hire or fire them.

Secondly, few persons suffering severe psychological distress are capable of functioning as *informed* "consumers", mostly because they are poor and some because of reduced competency. In a microeconomic social welfare theory which gives some importance to values of equity or fairness, a market might be theorized to provide a justifiable means of social allocation *if* players start off roughly equal and *if* the system does not contain a dynamic which permanently marginalizes classes of players, preventing them from rejoining the game.

Thirdly, only a fraction of a mental health system, if that, could even aspire to functioning as a competitive market, given the variety of assumptions which must be met that define such a market as efficient, including the requirement that the "good" bought and sold in the market — in effect, health — be a private good.

It is not inconceivable that elements of market behaviour could improve aspects of the mental health system, thereby advancing user needs (pursuant to an understanding of user needs as that which users would seek to satisfy in a market able to satisfy those needs), but there is little or no semblance of such a market in any mental health system today, and construction of such a market must ensure that basic microeconomic assumptions ensuring efficiency and fairness are adequately met (rather than confusing, as the public's implicit faith in our current economic system too often seems to do, a *laissez-faire* survival of the fittest model with the concept of a competitive market).

Best interests model. I believe that this has been the most influential, implicitly or explicitly, of all narratives underlying confidence in the mental health system, in that the system is viewed as doing what is "best" for patients — what is best for them being in effect what they "need", regardless of what they say they want.

Such a model cannot be seen as applicable to *competent* patients; there is no reason to think that anyone can decide for a patient what is best for them better than the patient could — unless, for example, one wishes to suggest the unquestioned authority of one class of persons to decide what is good for others. While priests and emperors are no longer held to possess such authority, some critical writers have suggested that society has accorded this authority to psychiatrists, with respect to those whom psychiatrists diagnose as mentally ill. Such an authority, if it is exercised on behalf of persons not considered incompetent, would then pertain to a medical meritocratic society, and would be more properly viewed as a medical model narrative than a best interests narrative.

However, insofar as mental patients are, by definition, implicitly or explicitly viewed as incompetent (e.g., under a "broken brain" theory), a best interests narrative might still explain a wide-spread confidence in the mental health system as meeting all or most patient needs. Such a view, however, holds an extremely tenuous relationship to the facts, when competence is tested by a functional measure of capacities which is *independent* of diagnosis and treatment refusal. While some studies have shown a limited correlation between some diagnoses and given competence tests, there is no evidence that psychiatric patients are substantially less competent than others.

Hence I would hold that the best interests narrative can only be a starter for psychiatric patients judged incompetent by an independent, objective test of functional incompetence — which test should meet adequate scientific criteria and have threshold levels for incompetence that are justifiable with respect to the area of decision making concerned.

With respect to incompetent patients, this thesis has found that the best interests model does not provide an objective for decisions on behalf of those patients that can be related in any scientifically demonstrable way to patient needs — expressed as "underlying" preferences or otherwise. It is not clear what the best interests objective tries to do, nor even that it is centered upon the patient rather than the interests of those deciding for the patient.

If patient need satisfaction is operationalized by trying to make the decision that the patient would make, as the "pure best interests" model developed in this thesis calls for, we find that there are a variety of considerations which should be incorporated in proxy processes but which apparently have not been.

Indeed, proxy decisions on behalf of psychiatric patients may systematically be biased in favour of medical or other perspectives and interests unrelated to the patients, inadequately account for the patient's own values and living conditions, do not account for the patient's attitude to risk and preference for controllable risk, and fail to consider the subjective impacts of forced treatment. The impact of such biases may be to systematically favour medical treatments which the patient has refused, despite underlying preferences which might accord with such refusal.

The literature which assesses the accuracy of proxy decision making, none of which concerns psychiatric patients, suggests that success in meeting underlying patient preferences is fair to poor. Given the above-mentioned problems, there is every reason to suspect that the track record with respect to psychiatric patients is worse. Proxy decision making is fraught with uncertainty from many sources; given the widely acknowledged importance in our society of autonomy values, a prudent rule of thumb in the face of such uncertainty is to accept the decision expressed by the patient, unless there is at least reasonable certainty that an estimated preference otherwise is accurate.

If it is true that psychiatric treatments have at best rather limited benefits over costs, as I believe, an appropriate policy with respect to proxy decision making on behalf of psychiatric patients would simply prohibit over-ruling expressed patient preferences with regard to such treatment decisions.

In sum, the best interests model is inapplicable to competent patients and has markedly failed to demonstrate that the interests of psychiatric patients designated as incompetent are adequately met. This would seem to hold true under any common understanding of patient interests, e.g., patient "needs", welfare, or underlying preferences, since there is no clear operationalization of any of these concepts under the best interests model.

The pure best interests model developed in this thesis aims to correct this deficiency, by clearly defining the objective of substituted decision making as making the decision that the patient would have made if temporarily competent, and developing criteria for evaluating proxy processes according to this objective.

Power Access Condition. This thesis outlined numerous sources of systematic power disadvantages for users compared to other actors such as family members, psychiatrists and other practitioners, care and treatment institutions and administrators, as well as their unionized workers, and of probably much more importance than is commonly recognized, pharmaceutical companies.

This thesis suggests, firstly, that users have little power as seen by *direct measures*, such as actual observed influence in mental health policy and politics. This is seen in case studies of the development of major policy initiatives, as described in the Deinstitutionalization and Right to Treatment papers.

Secondly, *outcome measures* also strongly suggest that the system does not reflect user power. Given the revendications of user groups, one would expect mental health policy and structures and ensuing therapeutic practice to be markedly different if they had been substantially influenced by the exercise of user power. The mental health system would be far less coercive, would be based around a community service model providing a wide variety of psychosocial services and resources (including housing and income), the role of institutional, medical and especially pharmaceutical care would be far less, users would be provided alternatives and allowed to make *informed* choices, and, especially, users would be prominent in the administrative and policy-making structures of mental health related institutions.

Existing user groups are few, poorly financed and almost invisible, but what they call for should provide a more reliable indication of user demands than the "satisfaction" surveys of health providers and the claims of lobby groups which pretend to represent users and their interests, but which are in fact dominated by family members, volunteers and practitioners.

What is rather more troubling, however, especially with respect to ethical and democracy theory criteria for a system which operates fairly and efficiently, is that, given the variety of power disabilities described in the previous part, *users could not be expected to have meaningful power in the system.* Yet users' lack of power in the mental health system and in therapeutic encounters, when it is recognized, is probably attributed by most to the disabling effects of their psychological distress, in that their supposed irrationality, incoherence, passivity, etc., render them *incompetent* to exercise power — hence the implicit or explicit conclusion that *of course* mental patients do not have power, nor could they or should they, until they have gotten "better" consequent to treatment beneficently "suggested" to them.

Such a view is problematic for several reasons, the most important being that if *lack* of power is a factor contributing to psychological distress, or to its severity as experienced by users, a system aiming to help users by imposing power upon them may be doing the opposite of what is needed.

Another reason is that insofar as psychological distress is related to incompetence — which it is not necessarily — "incompetence" has to be understood as pertinent to specific types of tasks, perceptions, reasoning abilities, problems, etc. Empowering a user who has impaired competence — i.e., facilitating the development of skills and confidence to overcome or ameliorate psychological distress or alter circumstances that contribute to such distress — can hardly be aided by further infantilizing that user — i.e., attributing incompetence across the board and hence making too many decisions for the user.

Rather, empowerment would more logically imply letting the user take some chances by experimenting with the exercise of power and learning from it, and working with the user's strengths so that as much as possible the user him or herself identifies and addresses those skills which need improvement.

The Policy Intervention Condition *(The Bureaucratic Model).* The Deinstitutionalization paper examined in detail the deinstitutionalization movement as a major reform effort which has gone awry, compared with expectations. That article illustrates how even well-intentioned "rational" policy, ostensibly favouring users, can be distorted in a complex program field structured by distributions of power, incentives, and constraints. With respect to the policy intervention condition, this shows the relative inefficacy of "government intervention" to change a system already characterized by power and interest divergences.

The Right to Treatment paper shows how well-intentioned reforms in the mental health system that do not address the system as a whole will frequently backfire. It concludes with a discussion of the ethical rationale for intervening in the lives of vulnerable persons, pursuant to the view elaborated in that paper and elsewhere in this thesis that the interests of patients will not be protected by relying on *paternalistic* concern for patient welfare, as such paternalism has been understood in the past.

Can we expect intervention at the administrative and service delivery levels to rectify the degree to which the system might otherwise not meet user needs — i.e., through inadequately meeting the interest convergence and power access conditions? The simple answer is that such would be impossible, since any intervention is instigated and implemented by human actors and groups of actors; insofar as the interests of those actors are divergent from those of users, and insofar as they are not themselves users, or pushed by users having power, it is hard to imagine how and why they would try to base the system more upon user needs.

While the discussion of interest convergence suggests that the market and medical model narratives clearly fail to support that condition, the best interests model remains ambiguous; it may be that many actors, including family members, some practitioners, and administrators have some strong incentives to *try* to meet the needs of users, independent of their other interests which conflict with those of users.

That possibility cannot be ruled out, but nor can it be found credible enough to provide substantial confidence that the mental health system will meet user needs — paternalistic beneficence, as intended by the best interests model, does not have a conception of "needs" which is independent of how needs are assessed and of who, and whose values and perspectives, do the assessing.

Policy Makers and Administrators: Constraints, Incentives and Interests. However, it may be too hasty to reject this condition on the basis of too-large generalizations about interest divergence and power disparities. Some actors who might be influential upon policy, at least at more local levels, may have interests that more closely converge with users than those focused upon in the discussion of interest convergence. It is not easy to identify such actors as a large class.

Psychiatrists. For example, while psychiatrists as a group appear to have interests strongly favouring a medical approach, which this thesis finds very unsatisfactory as a narrative that user needs are met, we might find in some jurisdictions that the nature of the profession, its organization, its financing, its government-sanctioned role in social control, and its history is such that psychiatrists in those jurisdictions have incentives and constraints which lead them toward a different set of guiding objectives for the practice of their profession.

It would be interesting to investigate why in some countries the role of the medical model is far less for psychiatrists, who may still associate with branches of psychiatric ideology that briefly flourished several decades ago: "community psychiatry", "radical psychiatry", "orthopsychiatry", etc. However, such exceptions are prominent only because of how completely biomedical the psychiatric profession has become in the western world.

Social work and nursing. The literature of the social work and nursing professions incorporates a substantial biopsychosocial perspective; however, while much practice, or at least discourse, of these professions had been, during the 1960s and 1970s, a reflection of that perspective, it seems that these professions, like the other helping professions, have concentrated in recent years in carving out a niche within a biomedical orientation. There is no shortage of literature, for example, critiquing clinical social work for its increasing biomedicalization and social control roles in conjunction with psychiatry. Social work training now typically includes the DSM, and psychiatric labeling has been accepted by much of the profession in helping to categorize their clients and their problems[35].

---

[35] However, it is noteworthy that the main and most extensive critique of the DSM approach has been authored by two social workers, Kirk and Kutchins (1994).

Professional administrators. Professional administrators and policy makers tend to adopt the language of community care, empowerment, and the biopsychosocial perspective. While there is no doubt that for many this language reflects a sincere effort to center upon user needs and empowerment, independent of a biomedical orientation and institutional and psychiatric dominance, it is difficult for such persons to resist the demands of powerful stakeholders, including the existing institutions and their representatives in the mental health systems for which they have responsibility.

Nevertheless, there are more than a few cases of dedicated individuals and mental health policy groups marginally moving the system toward being more user centered and less biomedical and institution-based. Such movement is aided by government priorities to reduce hospital and asylum-based spending, and also by user advocacy which is slowly growing more effective.

The former pressure can provide mixed blessings, however, as detailed in the Deinstitutionalization paper: reduced spending upon hospitals and asylums (actually more a question of curbing growth in such expenditures) has not translated into greater spending upon community-based services. Although community care has been found to be of equal or greater cost-efficacy than a hospital/asylum based system, the benefits take longer to make themselves felt, and an adequate community care system requires initial investments[36].

Hence the reality of community care systems is to be found more on paper than in the lived experience of those suffering psychological distress, who are faced with a variety of administrative structures, individualized service plans, and medication prescription as outpatients, but with few reasonable alternatives with respect to housing, work, education, accompanying and advocacy, counseling, etc.

With respect to policies explicitly aiming at user empowerment and participation or control in governing and policy structures, it is hard to imagine

---

[36] Yet certainly less than if the movement had been away from community care toward asylums and hospital. The current situation contains an irony in that a great deal of costs associated with asylums and hospitals are already "sunk" — they cannot be recovered — and maintaining such institutions over the short run may cost less than the investments required for community care. As long as policy retains a short-run perspective, it can be very difficult to make the transition to a system which over the long-run could be much cheaper.

them being meaningfully implemented without users themselves asserting themselves as important stakeholders, *a priori*.

Otherwise, we have seen the results of empowerment and participation policies as coopting users by focusing "empowerment" upon the mental health deficits of users — requiring pharmaceutical treatment and control by others as "enabling" empowerment — and participation policies which obscure the fact that users themselves are given a minimally nominal role, if that. "Community groups" and "community resources" may obtain influence in governing structures, but these groups ordinarily represent mostly professional (government paid) caregivers and, to a lesser extent, volunteer and family caregivers. In this context, if users themselves are consulted, it is usually in the form of "user satisfaction surveys", which measure degree of satisfaction with *existing* services.

If policy makers are able to make even marginal changes in the distribution of power *within* a service system which facilitate user advocacy, e.g., financing for user groups, membership in governing structures, programs to aid users in the representation of their own interests at the therapeutic, legal and policy levels, such an incremental change might have a chance to succeed in the long run, in that users who were formerly not seriously considered stakeholders by bureaucracies can become part of the constituency upon which they are dependent for credibility and survival. Under some circumstances, policy makers might succeed in advancing such marginal changes despite opposition of important stakeholders to the objectives of the reform[37].

As chaos theory demonstrates, very small marginal changes can have very major impacts. In social systems, distributions of power resources are of crucial importance for outcomes, especially over the long-run, since the more power one has the easier it is to increase it even further through further alteration of the rules of the game. Hence reform-minded policy makers should never give in to

---

[37] E.g., when those stakeholders are poorly organized; when the policy proposals are obscure, intentionally or otherwise, and/or are advanced in a confusing policy environment; when they are slipped by in the wake of other issues which attract more attention; when the result of concerted efforts by dedicated individuals with strong powers of persuasion; when sold as, in effect, providing token power to improve the image of service providers but the consequences are more than expected by policy proposers or approvers.

disillusion, but they should realize that perhaps the most effective way to advance reform that has been continually stymied is to target the sources of power — including but not restricted to capital — which sustain those whose interests hinder reform.

Alternatively, one might target interests: for example, would federal government prohibition of pharmaceutical company funding of university research, lobby groups, and promotional or "educational" payments to psychiatric associations, compensated by government funding, change the incentives of those actors in a way which would facilitate reduction of dependence upon psychoactive drugs?

On the other hand, too many simplistic applications of systems theories ignore the importance of supra-systems. A small burning candle may be the most visible object in a closed room, *until* sunlight pours into the windows. Similarly, particular service delivery structures operate in a much larger environment that remains largely untouched by changes at the local service level. Reforms at that level may be difficult to sustain given those environmental pressures. Local reform successes might be interpreted as failures, for example, by state-wide lobby groups with the means to finance ingenious "public information" campaigns, who fear that the local initiatives might spread throughout the state.

Systematic Failure of User-Centered Reforms

The discussion thus far has brought us to the following point: there is no current basis for an expectation that user needs will be met on the grounds of user access to power, nor upon interest convergence grounded on the medical or market models. A hypothesis that user needs are met involving such theory can, in fact, be firmly rejected. The best interests model alone is also insufficient for this expectation for two major reasons: the most influential actors have interests which significantly diverge from those of patients, and the best interests model has a very ambiguous relation to user needs, such that verification that they are met is difficult or impossible.

However, as discussed above, there may be sufficient convergence of the interests of users with some actors involved with policy-making and administration such that incremental relatively localized reforms might take place that have a chance of altering power structures in favour of users and even instituting a change dynamic that can permanently alter the system. While this is possible, it is likely to be restricted to the local level and, furthermore, might be subject to long-term environmental pressures that make it impossible to sustain.

This thesis has engaged in a great deal of generalization about mental health systems in order to characterize the most fundamental dynamics, actors, constraints and incentives that determine the general form of mental health systems over time. Nevertheless, there is wide variety of mental health systems throughout the different states of the U.S.A., for example. Careful comparative analysis will reveal some that show consistent, if usually slow, progress toward reform which significantly empowers users and gradually replaces an institutional biomedical system with a community-based psychosocial system. It would be pertinent to the objectives addressed in this thesis to carry out such a comparative analysis in order to isolate what factors and dynamics have most facilitated or stymied reform initiatives.

The generalization remains, however, as pointed out in the problématique, including the Deinstitutionalization paper, reinforced in the Right to Treatment case study, and analyzed in detail as a policy research problem in the Strategic Reform paper: that major user-centered reforms have failed completely or have fallen far short of expectations.

Such reform initiatives often *do* enormously change how the system operates and how it looks, and inevitably they have changed the distributions of incentives, constraints, and power (although for the most part the beneficiaries of such reforms are, as would be expected in a political economy analysis of changing social structures, among those who already had power; just the pecking order of the winners is somewhat shuffled).

Hence an asylum-based system has transinstitutionalized various classes of vulnerable or deviant persons to rest homes, semi-supervised rooming houses that are like asylums but smaller, prisons, institutions for mentally retarded persons, hospitals, etc., and placed "in the community" others who remain poorly integrated within social networks — despite the fact that many are lodged in the homes of their families — but whose deviancy is controlled by medications, outpatient commitment orders, and periodic visits to jails, hospital emergency wards, and chronic asylum stays.

Hence the "right to treatment" has succeeded in providing food and living conditions in asylums equivalent to prison inmates, but not in providing treatment and care alternatives to asylum inmates nor in facilitating their entry into the community and preventing their return to the asylum. Furthermore, this legal right, which seemed so promising pursuant to Judge Bazelon's decisions a quarter century ago, has become the right to the treatment decided by "professional judgment", according to the U.S. Supreme Court, and the right to *forced* treatment according to the discourse of family lobby groups.

Hence ostensible efforts to bring democracy and choice to mental health services, e.g., through regional and local governing bodies which include in their boards of directors a variety of "community stakeholders", may have achieved the *opposite* of what they may seem to have promised to users: such bodies remove *political* responsibility for mental health and related services, exercised through government departments reporting to cabinet ministers concerned with the ballot box, and place responsibility in the hands of people who are for the most part, directly or indirectly, *employees of major health and service institutions, often of hospitals*.

This has had the effect of placing reform initiatives in the hands of those with the *least* incentives to pursue them. From a political science standpoint, such structures are disastrous, in that they place State intervention in the hands of State employees (directly or indirectly) who in effect are responsible to no one but themselves. Insofar as the governing bodies are legislated to require participation by those who are not involved in the traditional medical institutions, one might at least hope that some reforms, e.g., downsizing, decentralizing and service delivery "in the community" will become possible.

But those taking an empowerment perspective of user needs should view such initiatives as the result of an uneasy temporary alliance with those advancing such reforms due to their own interests, rather than over-identifying with such reform processes, falsely assuming that they are related to true empowerment or psychosocial based reforms.

As the case studies in this thesis demonstrate, the "unexpected consequences" and "perverse effects" of major reform initiatives in the mental health system are so important that it seems justified to see them as *main* outcomes, whether forecast or not, whether included in the original discourse about the reform objectives or not.

They should be seen as what happens when one alters institutional structures, incentives and constraints, such that most of those previously holding most power in the old system find ways of adapting to the new system.

Hence, while many applaud the development of "community care" (although there is a growing backlash which threatens to halt or even reverse momentum toward the community), as long as users themselves possess no power in reform processes and new structures, we might expect to find the most fundamental interests of the powerful players manifesting themselves, but in new ways.

Those who wish to control deviance will find ways to do so, if not in asylums, straight-jackets and ice-baths then in "total community" devices using neuroleptic drugs, outpatient treatment orders, periodic jailing, hospital or asylum confinement (or the threat of such), and income, housing and social services contingent upon permitting government agents of social control — today often social workers — to enter, interfere with, coerce and control the choices of daily living.

Those who wish to ensure that psychological distress will be viewed and treated as brain dysfunctions — whether family members, psychiatrists or insurance companies — find this easy to do in the context of offering "informed choices" to users, when the nature of the information to be provided depends on who has the money to provide it. Hence current pressure on Canadian policy makers to permit advertising of pharmaceuticals directly to the public. Even if much of health policy and mental health services could be wrestled away from medical practitioners, a reasonable fall-back plan for pharmaceutical companies would be to show users and their family members how taking a drug implies that they are not "crazy" and that they are blameless for their suffering.

Those who wish to ensure that users stay in a subservient position with respect to mental health policy and administration can find many ways to coopt them — e.g., through luring user leaders to well-paid government positions, acculturating them to the new environment and ensuring through usual bureaucratic practice that they remain responsible to major stakeholders (who are not users); through allowing them to sit in committee and board meetings and then humiliating them through the use of the specialized language of bureaucrats and service providers; through providing chronic funding to user groups that rests dependent on the funders' perceptions of how well the groups "collaborate" in "partnership" with others, or which subtly alters the interests of the activists in the group (e.g., through drug company funding).

Those who wish to distract attention away from major social problems of poverty, job insecurity, environmental toxicity, etc., will find ways to do so even in a supposed biopsychosocial model, by focusing on the psychological aspects of empowerment — and consequently the deficits of users' mental reasoning abilities; by concentrating etiological research upon genetic factors; and by intervening in coercive and moralistic ways upon the "behavioural risk factors" in users' lifestyles.

Hence those concerned with an empowerment perspective of psychological distress, and disturbed by the degree to which the marginalization of users is systematically reinforced by the operations of the mental health system, should insist that the basic objectives of reform proposals are carefully analyzed with respect to their ethical or democracy theory content.

Furthermore, the ways in which those objectives might be operationalized has to be very carefully scrutinized in order to avoid cooptation and hidden agendas. In my view, empowerment-inspired planners and stakeholders should not hide their basic standpoints from others as a calculus of compromise or strategy; when these fundamental objectives remain hidden it is too easy for those with power in the system to guide the ways reform is implemented, without having to explicitly refer to those objectives.

Hence, for example, if the purpose of decentralization is supposed to be to increase responsibility at the local level, it is pertinent to ask "Responsibility to whom, *exactly*?" If the purpose of individualized service plans is supposed to be centering care upon the user as an individual person in unique contexts, it is pertinent to ask what role the user has in identifying needs, to what degree needs are identified according to the services available, and what happens if the user is dissatisfied with the plan. If the purpose of deinstitutionalization is to favour user autonomy and dignity, it is pertinent to ask to what degree each aspect of "community care" impacts on autonomy and dignity.

Directions for Future Research

As noted earlier in this thesis, its main aim was to help pose "better" questions. A very large number of questions arose or are implied by the foregoing discussions. Further exploring some of them would have improved this thesis, sharpened its arguments, and clarified policy implications. Further exploring others, while important for various aspects of the problems considered here, would have distracted attention from the systematic analysis of narratives purporting that the mental health system will meet user needs. Some questions properly need to be addressed from a much broader perspective than that of "mental health" concerns. Finally, much of the research that this thesis points to must develop new empirical observations in neglected areas, requiring major, funded research programs. Below are some of what I consider the more pressing or interesting research deficits or questions that should be addressed, and which flow from the concerns, observations and speculations raised in this thesis.

- **Psychiatry/Medicine as Religion.** In any policy area, there should always be room for research and speculation which tests the limits of "serious" research, partly because it can help to reconceptualize in essential ways the phenomena in which we are interested. The suggestions by some thinkers, including Szasz, that psychiatrists have become the priests of postmodernity are not entirely tongue-in-cheek. A broad cultural and historical perspective would recognize that the transformations of our societies, in which belief systems have always been strong but have taken different forms both in the natures of the beliefs and the ways in which they become institutionalized in social structures, may well imbue in groups like scientists, teachers, medical practitioners, or government bureaucrats, an authority that verges on the mystical. I think it would be fruitful for advancing understanding of the role psychiatry plays, and is accorded, in contemporary society, to compare an "institution of psychiatry" (various definitions are possible) with our understandings of institutions of religion.

- **The Legal System.** There have been ample indications in this thesis that the legal system has not helped users of the mental health system; in effect, it is as if society had decided that the legal system is not pertinent to users, allowing paternalistic caregivers the authority (whether formally approved by judges or not) to decide what is good for them. This is despite the acknowledged fact that users have *rights,* like any citizen, *not just "needs".* Systematic policy development work is required in order to give user rights substance (McCubbin and Cohen, 1997b), with attention to issues like access to lawyers, the unavailability of expert testimony counter to that of a treating psychiatrist, the nature of evidence that is and should be considered pertinent in court decisions, and the accountability of lawyers and judges to the public, users, and researchers for decisions which until now have been effectively cloaked in secrecy (justifying such secrecy in terms of the users' "right to confidentiality" rather than in terms of the prerogatives of those with authority over users).

- **"Helping" Professions.** Are the helping professions irretrievably founded upon paternalistic principles? Is it a paradox to expect them to integrate within their cultures, bodies of knowledge, and practices principles of empowerment aimed to liberate and enhance the capacities of those they are supposed to help? Can this paradox be surpassed — e.g., by reconceptualizing and reforming institutions in whose hands we place the task of advancing society's objectives with respect to health and welfare?

- **Forced Treatment.** This thesis has addressed a number of issues and ethical problems which should be considered with respect to the policy problem of forced treatment. Those matters need to be placed in a thorough, systematic *evaluation* of a policy allowing forced treatment, also taking into consideration issues barely dealt with in this thesis, the efficacy of treatments usually forced and their iatrogenesis. An evaluation of a *policy* must recognize that no policy can be perfect, e.g., banning forced treatment might be justifiable even if some patients might have benefited from it — perhaps in order to prevent abuses and other harms that overall would be more serious, or because it is difficult to identify who would benefit and who would be hurt. Furthermore, arguments favouring forced treatments should be put to the *community* of users: what would users themselves want to happen in the event of their eventual incompetence? Individual treatment refusals are routinely discounted as the expression of a deranged mind; however, is there any reason to think that when user groups *collectively* decry forced treatment the view is the product of a deranged *collectivity*?

- **Pharmaceutical Industry.** I will not apologize for suggesting in very strong terms that the research and planning communities dealing with the mental health system have been inexcusably negligent by not examining the role of the pharmaceutical industry in shaping the mental health system in ways counter to our overt public policy objectives as well as our basic conceptions of fairness. This thesis relied exclusively upon already existing literature, virtually none of which places attention upon the pharmaceutical industry in the context of how our mental health system operates and is transforming. Hence I have not been able to say as much in this regard as I suspect should be said. A community-wide research program should begin characterizing the industry and tracing out how its operations and dynamics impact upon the mental health system. The detailed tableau which emerges should be concerned with both the macro (e.g., influence upon political parties and governments as donors, taxpayers, employers, etc.) and micro levels (e.g., influence upon the practices of individual psychiatrists through support to psychiatrists' associations, gifts, and funding of their own research). Such research does not have to be pursued in a blind "attack the big corporations" mentality; it should aim to help policy makers make social choices, always involving trade-offs, that rationally reflect the interests of the public over the long-run.

- **User "Needs".** Should the concept of "needs" be allowed to play a role in mental health services and planning? Can "needs" have any sense apart from the objectives with which needs are associated? How could any caregiver or caregiving system determine user needs without explicit attention to what those objectives are? If the objectives are not the users', as conveyed by them, how could the associated needs be those of users?

- **User Competence.** This thesis has argued that therapeutic objectives in the system should be based on empowerment principles — facilitating the development of users' own capacities to take control over their lives. If this is a worthwhile objective, interventions should be designed and evaluated with this objective explicitly in the forefront. Hence, for example, when faced with a person considered incompetent, emphasis should be placed on helping the person to make their own decisions or intervening in such a way as to favour the development of competence. Until now, the mental health system has blithely assumed that routinely substituting "good" decisions for the expressed preferences of persons judged incompetent is therapeutic and "liberating" once imposed treatments take effect and reduce symptoms. This assumption should be rigorously evaluated. In particular, it is important to investigate to what extent standard treatments, e.g., neuroleptics and electroshock, *impair* competence over the short or long-run, and to what extent the mere fact of *imposing* treatments impairs competence (e.g., through a process of infantilizing). Empowerment principles suggest that the emphasis in the mental health system should be on working with user strengths to *enhance* competence, rather than paternalistically compensating for user weaknesses. Existing research in the mental health and psychiatric fields has not yet reflected this orientation.

- **User Wants.** This thesis did not investigate systematically and thoroughly the ways in which user *wants* have been explicitly assessed by service agencies; a meta-analysis in this regard would be useful. It has been occasionally pointed out that agencies usually survey *satisfaction* with existing services, or user *preferences* among existing services — if they survey at all — rather than more open-ended surveys which ask users what they need to improve their lives, reduce their distress, etc. Insofar as psychological distress is both a consequence of and contributor to poor quality of life, it might be worthwhile to conduct a survey which painstakingly avoids references to mental illness, mental health services, psychological distress, etc., and simply ask how life might be improved and what resources might help in such improvement. It may well be that efforts to improve the quality of life of users — which efforts need not be defined by traditional conceptions of "mental health services" — would be more effective in reducing psychological distress than efforts directly *aimed* at such distress. This might be the case not only because of the importance of social and economic factors in the development and amelioration of psychological distress, but also insofar as effective "psychological" intervention *must be mediated and controlled by the object of intervention* if it is to be effective in terms of reducing distress and enhancing capacities. Hence, while facilitating access to social networks may have the *effect* of improving psychological welfare, it may be unnecessary or even harmful for the persons doing the facilitating to plan the intervention as a psychological intervention.

- **User Representation of their Own Interests.** The lack of power of users in representing their own interests, at both service delivery and collective levels, is probably the most important observation of this thesis, even if it is the most obvious. Given that this does not seem to be a concern of most mental health systems, I think it is necessary to document the lack of user power and the existing ways in which users express their own interests even more systematically and thoroughly, and especially empirically, than this thesis has been able to do. This is more pressing at the level of collective action than in the clinical situation, to which a number of researchers have already directed their attention. What groups are users in? What groups do they dominate or control, both in terms of numbers and influence? How do they obtain financing? Do some forms of financing have the effect of coopting user groups? How do they reach the ears of those they wish to influence? Have user groups reached a "watershed point" in some jurisdictions where they become self-sustaining and enter a dynamic of growth in terms of their resources, membership, and influence? The results of such research, which might be conceived as a form of anthropological study of politics, would usefully be compared to what we know about groups which represent the interests of other actors in the system, including family members, caregivers and professionals, pharmaceutical companies, etc., in order to assess more clearly the relative powerlessness of users and the barriers they face in representing their own interests. The small literature on interest group politics in the mental health system has tended to group users together with their (usually temporary) allies when describing how policy debates unfolded; we need to gain a clearer idea of the nature and influence of user political advocacy as separate from the advocacy of those with interests convergent with users on specific issues.

- **Administrative Structures.** Many jurisdictions have "democratized" or "decentralized" the regulation and planning of mental and other health and social services. The rhetoric accompanying these "reforms" needs to be soberly assessed in terms of their supposed objectives. Indeed, it is not even clear that the objectives of these reforms include user participation in policy and administration. Do governing bodies include user participation? Is user participation token or important enough to be influential or determinative? Are those assumed to represent user interests truly representative of users and their interests? Are users truly represented by the "community resources" and "community participation" in regulatory boards? Do governing bodies act in such a way, e.g., pursuant to professional cultures, as to coopt or silence user representatives?

- **"Community Care".** This thesis provides a generalized description of modern mental health systems which is inadequate in many ways, partly a result of a literature which has not sufficiently reflected upon what post deinstitutionalization mental health systems do, who they serve, who does the serving, and, especially, how care and support systems are experienced on a daily basis by users. Existing formal mental health services need to be described in a context which includes formal and informal interventions beyond those falling under the aegis of "mental health". More generally, we need to better understand the lived experiences of users today, with reference to important social, economic and cultural dimensions of their lives. Users need to be understood not only in clinical or epidemiological terms but also as complex human beings living in particular circumstances. We need to be able to assess what "community care" actually means for users with respect to their own values pertaining to issues like dignity, autonomy, social support and integration, and quality of life; formal descriptions and evaluations of mental health services are not enough. In what sense have mental patients returned to the community? What does it mean for users to return to and live with "their families"? Might Goffman's (1961) analysis of the "total institution" be adapted to an understanding of an asylum without walls as the "total community"? Are the places where users live, including rooming houses and halfway homes, different in more than size from large rural asylums?

- **Long-Term Planning.** Reform in the mental health system has fallen prey to the same problems as in every other area of pressing societal concern: the inability of our systems of social choice — including but not limited to governments — to think and plan for the long-term, to invest accordingly, and to continually revise reform strategies in accordance with coherent long-term objectives. I think that this thesis has demonstrated the seriousness of this problem in the mental health system. Policy planners in the mental health as well as other policy areas would probably do well to incorporate in their problem definitions the basic structures of social choice. The ways in which national democracies are structured and govern, for example, should be of crucial concern to those planners, not just to specialists in the fields of elections, governments and bureaucracies.

Final Thoughts: The Rationality of the Mental Health System

With respect to the *interest convergence condition,* the thesis firmly rejects medical and market model narratives suggesting that through convergence of interests under those models the needs of users are or could be met. The best interests narrative is inapplicable to competent persons. With respect to incompetent persons, since it is unevaluable it cannot be rejected, but nor can it be accepted. Hence this thesis fails to accept the proposition that interests of powerful actors converge enough with those of users that user needs will be met.

With respect to the *power access condition,* it is obvious that users have virtually no power within the mental health system, hence it can be unambiguously rejected.

Insofar as either powerful actors in the system were to have interests convergent with users, or users themselves had power, policy intervention at the administrative and service delivery levels could be effective in further advancing patient interests; otherwise interventions could not be expected to accomplish major long-lasting user-centered reform of the system. However, given the conclusions failing to accept the interest convergence condition and rejecting the power access condition, the thesis fails to accept the *policy intervention condition.*

Given the rejection of the power access condition and the failure to accept the interest convergence and policy intervention conditions, this thesis concludes that there are no reasonable grounds to accept a hypothesis that the mental health system meets user needs. Hence, if the mental health system is rational — structured and acting on the basis of demonstrable *reasons* as a consequence of individual and collective choices bearing on the system's *purpose* — its rationality does not lie in meeting patient needs.

This thesis does not *prove* that patient needs are *not* met: an impossible task without some *a priori* agreed-upon definition of what patient "needs" are or what the concept of "needs" means. What I do find is that there is no convincing theory or narrative describing how needs — and, for that matter, interests — are or could be met, especially given that users have very little power at the therapeutic and policy levels.

If users could achieve "consumer sovereignty" at the therapeutic level, and become "powerful stakeholders" at the policy level, the policy and ethical concerns would be greatly reduced: their personal and political power could alter the incentives of actors whose interests would otherwise diverge from theirs — caregivers, administrators, policy makers, the public and politicians would be more responsive to user claims and sensitive to how the life experiences and values of users could be translated into a meaningful participation in the social, cultural, economic and political life of the institutions of society.

The implications of the discussion of power and empowerment earlier in this part are similar to those more often raised with respect to feminist, race, and poverty concerns, which highlight the intimate interconnections between the personal and the political. That connection is particularly poignant with respect to the mental health sector, in that it is possible to theorize, as many critical writers have, that "mental illness" is at least partly a consequence of systematic marginalization in society as an exercise of superior power — or, indeed, a label for a class of powerless persons when no other label seems more pertinent.

If either is the case, it has to be expected that *treatment* for mental illness — treatment understood in the broad sense incorporating the expected roles of mental patients — will not only accept powerlessness as a given, but will inevitably *reinforce* it.

Hence, we might have expected the failure of reforms which would have the effect of increasing user autonomy, provide informed choice with respect to therapeutic and service issues, which would enhance the abilities of users individually and collectively to pursue and advance their interests in their own lives, in their communities, in the mental health system, and in society at large.

Such reform failures are not merely the result of technical failures, nor of the difficulties recognized in the policy sciences of "engineering" social reform by rational planning, nor of particular "unforeseeable" events that pop up in history (such as the "fiscal crisis of the State"). These reform failures are *inevitable* in a mental health system whose actual function and purpose — whatever they may be — is evidently at odds with the prominent *discourse* that the system is designed to meet the "needs" of mentally sick individuals.

I would suggest that this situation creates a society-wide cognitive dissonance so strong that it is quite understandable to be finding the "right (of the patient) to treatment" actually meaning the right (of society) to *force* treatment (upon patients). In this context it is not overly colourful to remind readers of George Orwell's fictional classics about totalitarianism, which describe what happens when reality conflicts with discourse too strongly.

What is immediately of interest to a critical analyst of the mental health system is how studiously the role of users is ignored in the policy and politics literature — or, perhaps more accurately, it is so banal or irrelevant to note the lack of user power and the nature of power deficits that the issue fails to seep into the academic consciousness.

I suggest that much could be revealed about the nature of the mental health system, viewed from a wide perspective, by studying the academic discourse *about* the system as a political system. It would be interesting to test the hypothesis that academic analysts share the sanism which critical analysts decry with respect to mental health policy planners, administrators, judges, and society at large (Birnbaum, Perlin). If this were so, it would represent a remarkable failure in the critical objectivity required for scientific study of the system, and cannot help but impede the development of innovative policy reforms aimed at satisfying user needs or interests.

It would also be interesting to test the hypothesis that academics have become far too dependent upon the system, e.g., via research contracts and partnerships with service providers, governments, associations and lobby groups, and pharmaceutical companies[38]. This would be less of a problem *if* the providers of funds did not all favour different visions of mental health and mental health care than users themselves.

---

[38] But who would fund such research?

As suggested in the Deinstitutionalization paper, it does appear that society's view of insanity and the appropriate place of persons suffering psychological distress has been importantly altered over this century by developing ethical and social thought incorporating humanistic conceptions based on the dignity of every person. How sad it would be if we were to allow growth in our collective values — our social conscience and consciousness — to be stymied by impersonal systems of power and influence over which we lose control through negligence or short-sighted self-interest.

References

Note that each article and manuscript in the thesis is followed by its own references.

Ackoff, R.L. (1974). The systems revolution. *Long Range Planning, 7,* 2-20.

Anon. (1994a). Law places 36-month limit on SSI, SSDI benefits to substance abusers, tightens SSDI requirements. *Hospital and Community Psychiatry, 45,* 1051.

Anon. (1994b). Inpatient, outpatient care trade dominant roles in mental health services between 1955 and 1990. *Hospital and Community Psychiatry, 45,* 1052.

Anon. (1994c). News briefs. Depression screening effective. *Hospital and Community Psychiatry, 45,* 289.

Appelbaum, P.S., and Schwartz, W.F. (1992). Minimizing the social cost of choosing treatment for the involuntarily hospitalized mentally-ill patient: A new approach to defining the patient's role. *Connecticut Law Review, 24,* 433-485.

Bachelard, G. (1969). *La formation de l'esprit scientifique: Contribution à une psychanalyse de la connaissance objective*. Paris: J. Vrin.

Bachrach, L.L. (1989). Deinstitutionalization: A semantic analysis. *Journal of Social Issues, 45,* 161-171.

Balter, M.B., and Uhlenhuth, E.H. (1992). Prescribing and use of benzodiazepines: An epidemiological perspective. *Journal of Psychoactive Drugs, 24(1),* 63-64.

Barham, P. (1986). *Schizophrenia and Human Value: Chronic Schizophrenia, Science and Society.* New York: Basil Blackwell.

Bates, R.H. (1993). A reply. *World Development, 21,* 1077-1081.

Bateson, G. (1972). *Steps to an Ecology of Mind.* New York: Ballantyne.

Beck, J. C., and Parry, J. W. (1992). Incompetence, treatment refusal, and hospitalization. *Bulletin of the American Academy of Psychiatry and Law, 20,* 261-267.

Becker, F.W. (1993). The politics of closing state mental hospitals: A case of increasing policy gridlock. *Community Mental Health Journal, 29,* 103-114.

Bierman, Jr., H. (1989). A behavioral decision theory paradox. *Behavioral Science, 34,* 286-288.

Black, B.L. (1986). Institutional context and strategy: A framework for the analysis of mental health policy. In W.R. Scott and B.L. Black (Eds.), *The Organization of Mental Health Services: Societal and Community Systems,* pp. 233-276. Beverly Hills: Sage.

Bloche, M.G., and Cournos, F. (1990). Mental health policy for the 1990s: Tinkering in the interstices. *Journal of Health Politics, Policy and Law, 15,* 387-411.

Boffey, P.M. (1984, 13 Sept.). Failure is found in the discharge of mentally ill. Psychiatric group says their release into communities is 'a major tragedy'. *New York Times*, pp. 1, 13.

Bosso, C.J. (1994). The contextual bases of problem definition. In D.A. Rochefort and R.W. Cobb (Eds.), *The Politics of Problem Definition: Shaping the Policy Agenda*, pp. 182-203. Lawrence, Kansas: University Press of Kansas.

Boudreau, F. (1986, March). From asylum to mental health: Is the patient any better off? *Canada's Mental Health*, pp. 16-18.

Boudreau, F. (1987). The making of mental health policy: The 1980s and the challenge of sanity in Québec and Ontario. *Canadian Journal of Community Mental Health, 6*, 27-47.

Boudreau, F., and Lambert, P. (1993a). Compulsory community treatment? I. Ontario stakeholders' responses to "helping those who won't help themselves". *Canadian Journal of Community Mental Health, 12*, 57-78.

Boudreau, F., and Lambert, P. (1993b). Compulsory community treatment? II. The collision of views and complexities involved: Is it "the best possible alternative"? *Canadian Journal of Community Mental Health, 12*, 79-96.

Bowler, T.D. (1981). *General Systems Thinking: Its Scope and Applicability*. New York: North Holland.

Breggin, P.R. (1990). Brain damage, dementia and persistent cognitive dysfunction associated with neuroleptic drugs: Evidence, etiology, implications. *Journal of Mind and Behavior, 11*, 425-463.

Breggin, P.R. (1991). *Toxic Psychiatry*. New York: St. Martin's Press.

Brown, P. (1984). The right to refuse treatment and the movement for mental health reform. *Journal of Health Politics, Policy and Law, 9*, 291-313.

Brown, P. (1985). *The Transfer of Care: Psychiatric Deinstitutionalization and its Aftermath*. Boston: Routledge & Kegan Paul.

Buchanan, J.M. (1972). Toward analysis of closed behavioral systems. In J.M. Buchanan and R.D. Tollison (Eds.), *Theory of Public Choice: Political Applications of Economics*. Ann Arbor: University of Michigan Press.

Bullock III, C.S., and Rodgers Jr., H.R. (1976). Impediments to policy evaluation: Perceptual distortion and agency loyalty. *Social Science Quarterly, 57*, 506-519.

Caillé, A. (1986). *Splendeurs et misères des sciences sociales*. Genève: Droz.

Castel, R., Castel, F., and Lovell, A. (1979). *La société psychiatrique avancée: Le modèle américain*. Paris: Bernard Grasset.

Chamberlin, J. (1990). The ex-patients' movement: Where we've been and where we're going. *Journal of Mind and Behavior, 11*, 323-336.

Cohen, D. (Ed.). (1990). Challenging the Therapeutic State: Critical Perspectives on Psychiatry and the Mental Health System. *Journal of Mind and Behavior, 11* (3 and 4).

Cohen, D. (1992, 13 oct.). Les "enfants de Duplessis", des prisonniers politiques? *Le Devoir,* B12.

Cohen, D. (1994a). Quelles sont les conséquences sociales et psychologiques en terme de qualité de vie des neuroleptiques et de leurs effets secondaires? In Conférence de consensus, *Stratégies therapeutiques à long terme dans les psychoses schizophréniques. Textes des experts*, pp. 149-184. Paris: Frison-Roche.

Cohen, D. (1994b) Neuroleptic drug treatment of schizophrenia: The state of the confusion. *Journal of Mind and Behavior, 15,* 139-156.

Cohen, D. (Ed.). (1994c). Challenging the Therapeutic State, Part Two: Further Disquisitions on the Mental Health System. *Journal of Mind and Behavior, 15* (1 and 2).

Cohen, D., and McCubbin, M. (1990). The political economy of tardive dyskinesia: Asymmetries in power and responsibility. *Journal of Mind and Behavior, 11,* 465-488.

Contractor, N.S., and Seibold, D.R. (1993). Theoretical frameworks for the study of structuring processes in group decision support systems: Adaptive structuration theory and self-organization systems theory. *Human Communication Research, 19,* 528-563.

Cornes, R., and Sandler, T. (1986). *The Theory of Externalities, Public Goods, and Club Goods.* Cambridge: Cambridge University Press.

Crozier, M., and Friedberg, E. (1977). *L'acteur et le système.* Paris: Seuil.

Dahl, R.A. (1961). *Who Governs? Democracy and Power in an American City.* New Haven, Connecticut: Yale University Press.

Dain, N. (1989). Critics and dissenters: Reflections on "anti-psychiatry" in the United States. *Journal of the History of the Behavioral Sciences, 25,* 3-25.

De Greene, K.B. (1990). The turbulent-field environment of sociotechnical systems: Beyond metaphor. *Behavioral Science, 35,* 49-59.

DeLeon, P. (1988). *Advice and Consent: The Development of the Policy Sciences.* New York: Russell Sage.

Etzioni, A. (1982). The fight against fraud and abuse: Analyzing constituent support. *Journal of Policy Analysis, 2,* 26-38.

Etzioni, A. (1988). *The Moral Dimension: Toward a New Economics.* New York: The Free Press.

Etzioni, A. (ed.). (1995). *Rights and the Common Good: The Communitarian Perspective.* New York: St. Martin's Press.

Evans, R.G., and Stoddart, G.L. (1990). Producing health, consuming health care. *Social Science and Medicine, 31,* 1347-1363.

Everett, B. (1994). Something is happening: The contemporary consumer and psychiatric survivor movement in historical context. *Journal of Mind and Behavior, 15,* 55-70.

Fineman, N. (1991). The social construction of noncompliance: A study of health care and social service providers in everyday practice. *Sociology of Health & Illness, 13,* 354-374.

Foucault, M. (1962). *Maladie mentale et psychologie.* Paris: Presses Universitaires de France.

Foucault, M. (1972). *Histoire de la folie à l'âge classique.* St. Amand (Cher), France: Gallimard.

Fox, R.W. (1978). *So Far Disordered in Mind: Insanity in California, 1870-1930.* Berkeley: University of California Press.

Frankford, D.M. (1994). Scientism and economism in the regulation of health care. *Journal of Health Politics, Policy and Law, 19,* 773-799.

Geller, J.L. (1994). Issues in American psychiatry reflected in remarks of APA Presidents, 1844-1994. *Hospital and Community Psychiatry, 45,* 993-1004.

Gigliotti, G., and Rubin, J. (1991). The right to refuse treatment: An application of the economic principles of decision-making under uncertainty. *International Journal of Law and Psychiatry, 14,* 405-416.

Goffman, E. (1961). *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates.* New York: Anchor.

Goldie, N., and Fredén, L. (1991). A 'crisis' of closure and openness: The present state of the Swedish mental health system in the light of a policy of sectorisation. *Social Science and Medicine, 32,* 499-506.

Goodwin, S. (1989). Community care for the mentally ill in England and Wales: Myths, assumptions and reality. *Journal of Social Policy, 18,* 27-52.

Grob, G. (1991). *From Asylum to Community: Mental Health Policy in Modern America.* Princeton, New Jersey: Princeton University Press.

Groff, G.K. (1976). System theory and organizational decision making. In J.D. White (Ed.), *General Systems Theorizing: An Assessment and Prospects for the Future,* pp. 92-94. Washington, D.C.: Society for General Systems Research.

Grossman, W.D., and Watt, K.E.F. (1992). Viability and sustainability of civilizations, corporations, institutions and ecological systems. *Systems Research, 9,* 3-41.

Harris, G.T., Hilton, N.Z., and Rice, M.E. (1993). Patients admitted to psychiatric hospital: Presenting problems and resolution at discharge. *Canadian Journal of Behavioural Science, 25,* 267-285.

Hastings, M.M. (1986). The long-term care puzzle and mental health policy. *American Behavioral Scientist, 30,* 143-173.

Hill, D. (1983). *The Politics of Schizophrenia.* Lanham, Maryland: University Press.

Hofferbert, R.I., and Schäfer, G.F. (1982). The application of general systems methodology to the comparative study of public policy. *International Journal of General Systems, 8,* 93-107.

Hollingsworth, E.J. (1992). Falling through the cracks: Care of the chronically mentally ill in the United States, Germany, and the United Kingdom. *Journal of Health Politics, Policy and Law, 17,* 899-928.

Holstein, J.A. (1993). *Court Ordered Insanity: Interpretive Practice and Involuntary Commitment.* New York: Walter de Gruyter.

Horgan, J. (1993). Eugenics revisited: Trends in behavioral genetics. *Scientific American, 268,* 122-128.

Houthakker, H.S. (1950). Revealed preference and the utility function. *Economica, 18,* 159-174.

Illich, I. (1977). *Limits to Medicine. Medical Nemesis: The Expropriation of Health.* New York: Penguin.

Ingleby, D. (1985). Mental health and social order. In S. Cohen and A. Scull (Eds.), *Social Control and the State* (pp. 141-188). Oxford: Basil Blackwell.

Isaac, R.J., and Armat, V.C. (1991). *Madness in the Streets: How Psychiatry and the Law Abandoned the Mentally Ill.* New York: Free Press.

Jennings, A. (1994). On being invisible in the mental health system. *Journal of Mental Health Administration, 21,* 374-387.

Jones, K. (1988). *Experience in Mental Health: Community Care and Social Policy.* London: Sage.

Kanter, A. S. (1989). Homeless but not helpless: Legal issues in the care of homeless people with mental illness. *Journal of Social Issues, 45,* 91-104.

Katz, E., Nitzan, S., and Rosenberg, J. (1990). Rent-seeking for pure public goods. *Public Choice, 65,* 49-60.

Kemp, D.R. (1991). Mental health crisis in California. *Journal of Mental Health Administration, 18,* 154-166.

Kenig, S. (1992). *Who Plays? Who Pays? Who Cares? A Case Study in Applied Sociology, Political Economy and the Community Mental Health Centers Movement.* Amityville, New York: Baywood.

Kiel, L.D. (1992). The nonlinear paradigm: Advancing paradigmatic progress in the policy sciences. *Systems Research, 9,* 27-42.

Kiesler, C.A. (1992). U.S. mental health policy: Doomed to fail. *American Psychologist, 47,* 1077-1082.

Kiesler, C.A., and Sibulkin, A.E. (1987). *Mental Hospitalization: Myths and Facts about a National Crisis.* London: Sage.

Kirk, S.A., and Kutchins, H. (1994). The myth of the reliability of DSM. *Journal of Mind and Behavior, 15,* 71-86.

Kleinman, D.L., and Cohen, L.J. (1991). The decontextualization of mental illness: The portrayal of work in psychiatric drug advertisements. *Social Science and Medicine, 32,* 867-874.

Kuroda, T. (1993). A power index for multistage and multiagent decision systems. *Behavioral Science, 38,* 255-272.

Laing, R.D. (1967). *The Politics of Experience.* New York: Pantheon.

Lamb, H.R. (1994). Only good news is politically correct. *Hospital and Community Psychiatry, 45,* 517.

Latané, B., Nowak, A., and Liu, J.H. (1994). Measuring emergent social phenomena: Dynamism, polarization, and clustering as order parameters of social systems. *Behavioral Science, 39,* 1-24.

Lolas, F. (1994). Editorial. Medical praxis: An interface between ethics, politics, and technology. *Social Science and Medicine, 39,* 1-5.

Lurigio, A. J., and Lewis, D. A. (1989). Worlds that fail: A longitudinal study of urban mental patients. *Journal of Social Issues, 45,* 79-90.

MacIntyre, A. (1973). Ideology, social science, and revolution. *Comparative Politics, 5,* 321-342.

Marmor, T.R., and Gill, K.C. (1989). The political and economic context of mental health care in the United States. *Journal of Health Politics, Policy and Law,* 14, 459-475.

McCubbin, M. (1994a). A teleological reinterpretation of mental health system mutations. *Cahiers du Ph.D.-SHA Doctorat en Sciences humaines appliquées, 94-02.* Montreal: Université de Montréal.

McCubbin, M. (1994b). Deinstitutionalization: The Illusion of Disillusion. *Journal of Mind and Behavior, 15,* 35-53.

McCubbin, M. (1994c). *A Systems Theory Perspective of Mental Health Services and Policy Formation.* Unpublished monograph.

McCubbin, M. (1997a). Error under a pure best interests model of proxy decision making: Implications for the justifiability of forced treatment. Unpublished manuscript, presented to the XXIInd International Congress on Law and Mental Health, Montréal, June.

McCubbin, M. (1997b). Population Health: A call for breadth (mental health) and depth (psychosocial theory). Unpublished manuscript.

McCubbin, M. (1997c). Free will, rationality, and explanation: Toward a teleological *human* science. Submitted to *Journal of Mind and Behavior,* 29 Nov. 1997.

McCubbin, M. (1997d). General System Theory as a postmodern epistemology for the social sciences. Unpublished manuscript.

McCubbin, M., and Cohen, D. (1996). Extremely unbalanced: Interest divergence and power disparities between clients and psychiatry. *International Journal of Law and Psychiatry, 19,* 1-25.

McCubbin, M., and Cohen, D. (1997a). Mental health policy from a systemic perspective: Orientations for strategic reform. Submitted to *Journal of Health Politics, Policy and Law,* 25 Nov. 1997.

McCubbin, M., and Cohen, D. (1997b). Les droits des usagers de services en santé mentale: Le noeud étroit du pouvoir, de la loi et de l'éthique. Submitted to *Santé mentale au Québec,* 14 Nov. 1997. Presented to *Colloque institutionalisation/désinstitutionalisation dans un virage fou,* Laval, 24 Oct. 1996.

McCubbin, M., and Weisstub, D.N. (1998a). Toward a pure best interests model of proxy decision making for incompetent psychiatric patients. *International Journal of Law and Psychiatry, 21,* 1-30.

McCubbin, M., and Weisstub, D.N. (1998b). "Meeting the needs of the mentally ill": A case study of the "right to treatment" as legal rights discourse in the U.S.A. Unpublished manuscript, abstract submitted to the XXIIIrd International Congress on Law and Mental Health.

McNeil, D.H. (1991). The 'problem' is the problem. In S.C. Holmberg and K. Samuelson (Eds.), *Systems Science in the 21st Century; Integrating the New Sciences of Complexity in Service of Humans and their Environment. Vol. I: Advances in Natural Systems Science. Advances in Informatics, Communications and Systems Methodology,* pp. 183-189. Östersund, Sweden: Östersund University.

Mechanic, D., and Rochefort, D.A. (1990). Deinstitutionalization: An appraisal of reform. *Annual Review of Sociology, 16,* 301-327.

Mercier, C. (1989). Le role des facteurs subjectifs dans la fidélité a la médication. *Revue canadienne de psychiatrie, 34,* 662-667.

Meyer, J.W. (1985). Institutional and organizational rationalization in the mental health system. *American Behavioral Scientist, 28,* 587-600.

Miedema, B. (1994). Control or treatment? Experiences of people who have been psychiatrically hospitalized in New Brunswick. *Canadian Journal of Community Mental Health, 13,* 111-122.

Mirowsky, J., & Ross, C.E. (1989). *Social Causes of Psychological Distress.* New York: Aldine de Gruyter.

Morrissey, J.P., Goldman, H.H., and Klerman, L.V. (1985). Cycles of institutional reform. In P. Brown (Ed.), *Mental Health Care and Social Policy,* pp. 70-98. Boston: Routledge & Kegan Paul.

Mueller, D. C. (1989). *Public Choice II.* Cambridge: Cambridge University Press.

Muth, J.F. (1961). Rational expectations and the theory of price movements. *Econometrica, 29,* 315-335.

Nerlove, M. (1958). Adaptive expectations and cobweb phenomena. *Quarterly Journal of Economics, 72,* 227-240.

Oaks, D. (1993). Clue: Government & industry & phony "consumer" groups unite. *Dendron, 33/34,* 20.

Ontario Friends of Schizophrenics (1993). *Schizophrenia: Compassion Through Understanding* (brochure). Toronto: OFOS.

Paldam, M. (1993). Towards anarchy in macroeconomics: A wild review essay. *European Journal of Political Economy, 9,* 447-457.

Parker, R. (1993, Spring). Can economists save economics? *American Prospect,* 148-160.

Parsons, T. (1951). Illness and the role of the physician. *American Journal of Orthopsychiatry, 21,* 452-460.

Parsons, T. (1974[1957]). The mental hospital as a type of organization. In Y. Hasenfeld and R.A. English (Eds.), *Human Service Organizations: A Book of Readings,* pp. 72-97. Ann Arbor: University of Michigan Press.

Perlin, M.L. (1991). Competency, deinstitutionalization, and homelessness: A story of marginalization. *Houston Law Review, 28,* 63-142.

Pettit, P. (1993). *The Common Mind: An Essay on Psychology, Society, and Politics.* New York: Oxford University Press.

Pilgrim, D. (1990). Competing histories of madness: Some implications for modern psychiatry. In R.P. Bentall (Ed.), *Reconstructing Schizophrenia,* pp. 211-233. London: Routledge.

Polanyi, M. (1967). *The Tacit Dimension.* Garden City, New York: Anchor.

Popper, K.R. (1979). *Objective Knowledge: An Evolutionary Approach.* Oxford: University Press.

Prior, L. (1991). Community versus hospital care: The crisis in psychiatric provision. *Social Science and Medicine, 32,* 483-489.

Rachlin, S. (1989). Rethinking the right to refuse treatment. *Psychiatric Annals, 19,* 213-222.

Radford, J.P., and Carter Park, D. (1993). "A convenient means of riddance": Institutionalization of people diagnosed as "mentally deficient" in Ontario, 1876-1934. *Health and Canadian Society, 1,* 369-392.

Rawls, J. (1971) *A Theory of Justice.* Cambridge: Belknap.

Reed, S.C., and Lewis, D.A. (1990, Spring-Summer). The negotiation of voluntary admission in Chicago's state mental hospitals. *Journal of Psychiatry & Law,* 137-163.

Regan, T.G. (1987). Negotiating an innovative mental health service: The question of limits to negotiated orders. In D. Coburn, C. D'Arcy, G.M. Torrance, and P. New (Eds.), *Health and Canadian Society: Sociological Perspectives,* pp. 501-514. Markham, Ontario: Fitzhenry & Whiteside.

Renaud, M. (1978). On the structural constraints to state intervention in health. In J. Ehrenreich (Ed.), *The Cultural Crisis of Modern Medicine,* pp. 101-120. New York: Monthly Review Press.

Richards, D. (1990). Is strategic decision making chaotic? *Behavioral Science, 35,* 219-232.

Rochefort, D.A. (1984). Origins of the "Third Psychiatric Revolution": The Community Mental Health Centers Act of 1963. *Journal of Health Politics, Policy and Law, 9,* 1-30.

Rochefort, D.A. (1993). *From Poorhouses to Homelessness: Policy Analysis and Mental Health Care.* Westport, Connecticut: Auburn House.

Rochefort, D.A., and Cobb, R.W. (1994). Problem definition: An emerging perspective. In D.A. Rochefort and R.W. Cobb (Eds.), *The Politics of Problem Definition: Shaping the Policy Agenda,* pp. 1-31. Lawrence, Kansas: University Press of Kansas.

Rochefort, D.A., and Portz, J.H. (1993). Different systems, shared challenges: Assessing Canadian mental health care from a U.S. perspective. *American Review of Canadian Studies* (Spring), 65-82.

Rose, N. (1986). Law, rights and psychiatry. In P. Miller and N. Rose (Eds.), *The Power of Psychiatry,* pp. 177-213. Cambridge, United Kingdom: Polity.

Rosen, R. (1974). Some temporal aspects of political change. *International Journal of General Systems, 1,* 93-103.

Rothman, D.J. (1980). *Conscience and Convenience: The Asylum and its Alternatives in Progressive America.* Boston: Scott, Foresman.

Rothman, D.J. (1993). A century of failure: Health care reform in America. *Journal of Health Politics, Policy and Law, 18,* 271-286.

Salem, D.A. (1990). Community-based services and resources: The significance of choice and diversity. *American Journal of Community Psychology, 18,* 909-915.

Samuels, W. J. (1989). The methodology of economics and the case for policy diffidence and restraint. *Review of Social Economy, 47,* 113-133.

Scheff, T.J. (Ed.) (1975). *Labeling Madness.* Eaglewood Cliffs, New Jersey: Prentice-Hall.

Scott, W.R. (1985). Systems within systems: The mental health sector. *American Behavioral Scientist, 28,* 601-618.

Scull, A. (1977). *Decarceration.* New York: Prentice Hall.

Scull, A. (1979). *Museums of Madness: The Social Organization of Insanity in Nineteenth-Century England.* London: Allen Lane.

Segal, J.Z. (1993). Strategies of influence in medical authorship. *Social Science and Medicine, 37,* 521-530.

Shadish, W.R., Jr. (1984). Policy research: Lessons from the implementation of deinstitutionalization. *American Psychologist, 39,* 725-738.

Shakun, M.F. (1981a). Formalizing conflict resolution in policy making. *International Journal of General Systems, 7,* 207-215.

Shakun, M.F. (1981b). Policy making and meaning as design of purposeful systems. *International Journal of General Systems, 7,* 235-251.

Simmons, H.G. (1990). *Unbalanced: Mental Health Policy in Ontario, 1930 - 1989.* Toronto: Thompson.

Simon, H.A. (1969[1956]). Rational choice and the structure of the environment. In F.E. Emery (Ed.), *Systems Thinking: Selected Readings,* pp. 214-229. Harmondsworth, Middlesex, England: Penguin.

Snow, R.M., and Bloom, A.J. (1992). Ethical decision making styles in the workplace: Underlying dimensions and their implications. *Systems Research, 9,* 35-45.

Spector, M., and Kitsuse, J.I. (1977). *Constructing Social Problems.* Menlo Park, California: Cummings.

Staples, L.H. (1993). *Consumer Empowerment in the Massachusetts Mental Health System: A Comparison of Attitudes, Perceptions, and Opinions Within and Between.* Ann Arbor, Michigan: UMI Dissertation Services.

Strauch, B. (1997, 10 Aug.). Use of antidepression medicine for young patients has soared: To bolster market, makers seek F.D.A. sanction. *New York Times, CXLVI,* pp. 1, 24.

Szasz, T. (1974). *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct* (revised edition). New York: Harper & Row.

Szasz, T. (1990). *Insanity: The Idea and Its Consequences.* Toronto: John Wiley.

Szasz, T. (1991). *Ideology and Insanity: Essays on the Psychiatric Dehumanization of Man.* Syracuse, New York: Syracuse University Press.

Szasz, T. (1994). *Cruel Compassion: Psychiatric Control of Society's Unwanted.* Toronto: John Wiley.

Tanaka, H. (1989). Power as maximizing behavior. *Behavioral Science, 34,* 199-206.

Thompson, J.W. (1994). Trends in the development of psychiatric services, 1844-1994. *Hospital and Community Psychiatry, 45,* 987-992.

Torrey, E.F. (1988). *Nowhere to Go: The Tragic Odyssey of the Homeless Mentally Ill.* New York: Harper and Row.

Tyhurst, J.S., Chalke, F.C.R., Lawson, F.S., McNeel, B.H., Roberts, C.A., Taylor, G.C., Weil, R.J., and Griffin, J.D. (1963). *More for the Mind: A Study of Psychiatric Services in Canada.* Toronto: Canadian Mental Health Association.

Varian, H.R. (1984). *Microeconomic Analysis* (2nd ed.). New York: W.W. Norton.

von Bertalanffy, L. (1981). *A Systems View of Man.* Boulder: Westview.

von Foerster, H. (1988). La construction d'une réalité. In P. Watzlawick (Ed.), *L'invention de la réalité: Contributions au constructivisme* (trans. A.-L. Hacker), pp. 45-69. Paris: Seuil.

Wakefield, J.C. (1992). Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder. *Psychological Review, 99,* 232-247.

Wald, P.M., and Friedman, P.R. (1978). The politics of mental health advocacy in the United States. *International Journal of Law and Psychiatry, 1,* 137-152.

Warner, R. (1985). *Recovery From Schizophrenia: Psychiatry and Political Economy.* London: Routledge & Kegan Paul.

Warner, R. (1989). Deinstitutionalization: How did we get where we are? *Journal of Social Issues, 45,* 17-30.

White, D. (1992). (De)-constructing continuity of care: The deinstitutionalization of support services for people with mental health problems. *Canadian Journal of Community Mental Health, 11,* 85-99.

White, D. (1993). Les processus de réforme et la structuration locale des systèmes. Le cas des réformes dans le domaine de la santé mentale au Québec. *Sociologie et sociétés, XXV,* 77-97.

White, D., and Mercier, C. (1991a). Coordinating community and public-institutional mental health services: Some unintended consequences. *Social Science and Medicine, 33,* 729-739.

White, D., and Mercier, C. (1991b). Reorienting mental health systems: The dynamics of policy and planning. *International Journal of Mental Health, 19*(4), 3-24.

Wilson, R. W. (1997). American political culture in comparative perspective. *Political Psychology, 18,* 483-502.

Windle, C. (1994). Mental health services research: The shift to research synthesis. *Administration and Policy in Mental Health, 21,* 263-267.

Wolin, S. (1969). Political theory as a vocation. *The American Political Science Review, 63,* 1062-1082.

Free Will, Rationality, and Explanation

Toward a Teleological *Human* Science

Michael McCubbin

Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec), Canada, H3C 3J7.

[redacted]

Submitted to *Journal of Mind and Behavior*, 29 Nov. 1997
(Revised 29 Jan. 1998)

Abstract—This article explores the consequences for the nature of the human sciences of accepting the assumption of free will. After justifying the holding of such an assumption, the paper observes that explicit acceptance of the free will assumption has implications for the human sciences whereby modes of explanation and prediction are altered: toward the use of explanation for prescription rather than for prediction.

A "final cause", or teleological, emphasis in explanation posits the importance of the objects of desire, as conceptualized or idealized, and mutating. Mated with strategic behaviour, we find humans taking paths which are non-linear (or apparently so). Behaviour which may be explicable, when mutating goals and self-correcting behaviour are considered, may be inexplicable or misunderstood, when explained in terms of propelling forces or "causal" factors. Hence the result that behaviour which appears irrational when determinism is implicitly assumed might appear rational when free will and goal-oriented behaviour is assumed.

This article argues further that free will necessitates rationality, the existence of purposeful behaviour implies rationality and free will, and purposeful behaviour can only be understood in the context of value. The ability of any social science to predict non-trivial outcomes is likely to be always so limited that it remains doubtful that simple prediction should be seen as a primary objective of the social sciences. By the very act of describing or predicting behaviour we substantially influence it. The values underlying what a scientist chooses to describe, and how to describe it, convey, explicitly or not, a statement about the values she holds, and therefore those that she will propose.

This has important consequences for the human sciences, placing them on an ontological footing quite different from that of the natural sciences, and justifying concern with human values and aims as the basic matter with which the human sciences should be conducted. We would have a better chance of identifying *future* rather than past constraints by evaluating the actual choice problems faced by the various parties, past and present, with respect to constraints as they would see them, and with respect to their own values. Would we not therefore be more useful, as social scientists, by openly and actively participating in society's formulation of values and objectives?

Such a science I call "teleological human science", to distinguish it from the "scientistic" social sciences modeled after deterministic conceptions of natural science which ignore human agency and pretend to be value-free.

Free Will, Rationality, and Explanation

Toward a Teleological *Human* Science

Introduction

This paper explores the consequences for the nature of the human sciences of accepting the assumption of free will. After justifying the holding of such an assumption, the paper observes that explicit acceptance of the free will assumption has implications for the human sciences whereby modes of explanation and prediction are altered: toward the use of explanation for prescription rather than for prediction.

A "final cause", or teleological, emphasis in explanation posits the importance of the objects of desire, as conceptualized or idealized, and mutating. Mated with strategic behaviour, we find humans taking paths which are non-linear (or apparently so). Behaviour which may be explicable, when mutating goals and self-correcting behaviour are considered, may be inexplicable or misunderstood, when explained in terms of propelling forces or "causal" factors.

Hence the result that behaviour which appears irrational when determinism is implicitly assumed might appear rational when free will, goal-oriented behaviour is assumed. This has important consequences for the human sciences, placing them on an ontological footing quite different from that of the natural sciences, and justifying concern with human values and aims as the basic matter with which the human sciences should be conducted.

Such a science I call "teleological human science", to distinguish it from the "scientistic" social sciences modeled after deterministic conceptions of natural science which ignore human agency and pretend to be value-free.

Free Will

Strategic Belief

Even if we do not *know* that we have free will, we can, for *strategic* reasons, *choose* to believe in it. If we believe that we do not have free will, then it is probable that our enjoyment of life, our health, etc., will be less satisfactory than for those who were destined to believe in free will. The feeling of powerlessness associated with such a belief could be expected to reduce happiness.

If we believe that we do not have free will and are wrong, we will have paid that cost in loss of happiness, whereas being right in this belief brings no benefits. Hence it is best to convince ourselves that we have free will, *even if we think we are likely to be wrong.* There is no contradiction in believing something that we think is probably wrong: "belief" is a psychic term that is not completely dependent upon subjective estimates of probabilities. Belief is an *action* of the human mind that enables or disables future action alternatives.

Hence, an improbable belief might reasonably be adopted if the expected benefits of holding it outweigh its costs (future stream of benefits adjusted by probability and preference for current versus later enjoyment) compared to the alternatives. This pragmatic approach to belief might provide the "reason for" holding beliefs for their coherence with other beliefs, as opposed to a rationale for belief-holding based on acceptance of its propositions[1].

It may be objected that the above confuses belief, a thought, a condition, with choice, an action, a process. However, there is no need to "believe" in something in order to choose actions consistent with that belief. This view correctly sees belief as a different order of reality from choice but misinterprets it. Rather than seeing belief as relatively fixed in one dimension, while various choices are made in other dimensions, belief should be seen as formed by choice and in a context of choice: in a context of choice in that previous choices affect the sets of options available to us; formed by choice in that we can

---

[1] Gardenfors (1992) discusses this with respect to "epistemic entrenchment".

imagine a level of abstraction in the thought process that structures lower level patterns of thought. Our "beliefs" are in a dialectic with our choices.

Myth. While perhaps it is conceivable that one could go through life "acting as if" one held certain beliefs without actually holding them, it is probably inefficient to do so; e.g., it might be difficult for the student with two years left in her doctoral program to fully benefit from the choice to continue it while "knowing" cancer will kill her within six months[2].

Hence, it may be "rational" — reasonable — to "fool yourself" by adopting beliefs despite their unlikelihood. In order to avoid cognitive dissonance, one may adopt a *myth* containing "facts", including subjective estimates of probabilities, that are altered in order to make the adopted belief appear consistent with the facts. According to Koons (1992),

> ... rational agents must engage in some sort of as-if reasoning: they must (for want of a better alternative) operate on the basis of deliberate misrepresentations of the situation that, unlike the accurate representation, are not afflicted by cognitive blindspots[3] (p. 152).

The myth provides a justification to others and to oneself. It is obvious that inrespectful daily communication, it is difficult to give a detailed exposition of one's rationales; this may also be true of communication internal to the individual mind. The thought processes structuring belief may not be easily accessible from the ordinary processes of conscious verbal thought.

The Essence of Free Will

None of the above is meant to imply that free will is unlikely, only that even if it were, it seems reasonable to adopt such a belief. I have not seen a modern discussion of free will that clearly describes a concept providing the essence of free will — as an essence that has conceptual independence from the causal process. Hacking (1992), von Wright (1974) and Munn (1960) relate free will to

---

[2] See Pearce, 1989: "Enmeshment: The Willing Suspension of Disbelief", pp. 71-72.

[3] Koons developed the argument with respect to belief paradoxes.

existence of alternatives, uncertainty and unpredictability. These arguments establish the existence or likelihood of indeterminacy or unpredictability, thereby establishing a necessary but not sufficient proof for free will.

Determinacy. This line of thought might seem to suggest the idea that if any free will exists at all, that will is freer the greater the number of alternatives and the lesser their predictability. However, what meaning can free will have if reality is totally indeterminate and unpredictable? If one could not make relatively reasonable predictions under various action scenarios there would be no point in investing time in decision-making processes, and of course if the consequences of one's actions had nothing to do with the actions, then free will is an empty concept.

Hence, we could say that some determinism, as well as some indeterminism, are necessary conditions for the existence of free will. I will call this condition "constrained indeterminacy". Therefore we might suspect that the *crucial* ingredient with respect to free will lies elsewhere than with the existence of alternatives.

This concept of constrained indeterminacy weds the immediate future with the immediate past as a branching of the past into the possible future[4]. My suspicion is that we are so wedded to our conceptions of time — e.g., unidirectional and linear — that we may be missing other possible ways to conceptualize phenomena. This is understandable since, pursuant to the incompleteness theorem, we can no more define the limits of consciousness than we can the limits of the universe[5]. We don't know if there are dimensions beyond the fourth, and we don't know how something as complex and mysterious as "intelligence" or "consciousness" — consisting of information rather than matter — interacts with those dimensions.

---

[4] For a discussion of this problem with respect to humanity in general, see Teilhard de Chardin, 1955, especially "Le Problème de l'Action", pp. 226-234.

[5] Von Bertalanffy, 1981, "The Mind-Body Problem: A New View", pp. 85-108.

Foucault recognized society's inability to understand madness once insanity was placed in a context of restrained liberty[6]. Similarly, how can we understand the human unless human choices have some freedom from determinants? For to understand human behaviour as a consequence of prior conditions is not to understand the human, but rather to *explicate* motions within a process of motions: it is to implicitly define the human with respect to what the human is not. One might argue for the scientific usefulness of such an approach if we view the human mind and its preconditions as belonging to the same order of reality; a reality "alpha" with discernible boundaries: in such a case the human could be understood as what that order of reality alpha consists of, less that reality "beta" which is not-human.

There are two problems with such an approach. If the assumption regarding the order of reality is correct, then what is defined as human is a merely arbitrary matter of nosology. If behaviour is understood as caused then the concept "human" has no independence from other concepts: e.g., the mathematical expression Action as a function of Human as a function of Prior-Conditions can be simplified to Action as a function of Prior-Conditions. There is no need for the Human in this analysis, which explicates actions while leaving definition of the Human as redundant. In this case, therefore, the term "human science" would also be redundant, since in this analysis an efficient science must eliminate the human.

Consciousness. It is commonly stated that the *essence* of the human, that which makes it unique from all other forms of sensible reality, is consciousness. At a rudimentary level, consciousness might be understood, as it is by some writers in the information theory field, as implying an ordering of elements within one system that can only be understood within another system (a "higher" system, in that the system provides an accounting for the "lower" system), the process of understanding being the manifestation of consciousness (Campbell, 1982). One cannot obtain meanings from a book by making inventory of the letters and their juxtapositions.

---

[6] Foucault, 1954, p. 100: "c'est que le monde, en aliénant sa liberté, ne peut reconnaître sa folie."

Consciousness orders the sensible world; it constructs and reconstructs reality and its narratives. It has a beingness that is not completely logically dependent upon the lower level reality that it orders, anymore than a "book" is conceptually dependent upon the quantities of letters and spaces it contains.

Therefore, if the essence of human is the consciousness which *orders* prior conditions, it is of a different order from that non-human reality referred to above as beta; it cannot be defined in terms of beta as that part of reality alpha that is not beta. Since there can be no inference to what is human from what is not, a science which purports to study the human by studying what is not-human would be perverse, guaranteed to fail in its stated purpose.

To exclude free will from the human sciences simply because we don't understand it would be as perverse as rejecting "gravity" from the natural sciences because our successive theories of gravity have been shown to be wrong (see Sciama, 1959). We observe objects falling down, as we observe the choices of individuals changing the trajectories of the future. The *natural* sciences created the concept of the "force" of gravity; the *human* sciences are better off creating the concept of the "force" of consciousness than replacing an abstraction from the human world with abstractions from the non-human world.

Cassirer (1944) well illustrated the problem of using mechanistic models to explain behaviour in the human world with his observation that *machines are built for a purpose*. Somebody had to build them, and the person who did may well have had a purpose in mind that cannot be completely inferred or defined by comparing the sensible inputs with the sensible outputs.

The Search for Rationality

Critiques of Rationality Principles in Scientific Explanation

In my view there are certain basic propositions which, as obscure, ambiguous, problematic and unquantifiable as they may be, provide coherence to each of the social sciences. Political science sees the person as a "political animal" who seeks empowerment within his environment (Hobbes, 1983[1651]). Sociology posits the person as a fundamentally "social being" who needs, or wants, to live in the company of others. Psychology seeks to understand the choices of humans as expressions of self-reflective and self-aware consciousness. Economics, until it became a sub-field of statistics and mathematics, saw the individual as seeking to gain the most satisfaction or happiness possible (Etzioni, 1988). These propositions were valuable because they created a research agenda: the best theory is not evidently "true", but fecund in creating further hypotheses and thereby advancing science.

The most basic task of science is to explain. No modern scientist in the natural world would be satisfied with finding that A causes B, the scientist must explain *why*. Such an explanation shows what forces were involved in the A to B processes, and what the important conditions seem to be for the A to B process. The "truth" in the natural sciences is elusive; the task of the natural scientist is in fact not to find the "truth" but through a dialectic of abstraction and experimentation gradually tease out that which conforms to the truth (Bachelard, 1969).

Historians and epistemologists of natural science, like Bachelard, Cassirer (1944), and Kuhn (1970) have noted a "rupture" or "revolution" in scientific thought this century whereby mechanistic, linear, simple causality models of nature are being replaced by complex, chaotic, relatively unpredictable visions of reality. How ironic it is, therefore, that "modern" social science is imitating outdated natural science models in order to appear "scientific" (Castoriadis, 1978). Not only are techniques from one field being inappropriately adapted for use in another, but the techniques are, with the exception of greater statistical sophistication[7], 50 to 100 years old.

---

7. The science of statistics has made enormous advances in recent decades. This may in fact have created more problems than benefits in the social sciences however. Now that data is easily manipulable in computers, and powerful statistical programs are easy to operate, too often statistical programs are used without adequate understanding of the underlying assumptions, and too often regressions and correlations are churned out without theory justifying them. An accomplished methodologist recognizes that statistics are tools for testing theory based hypotheses; otherwise they may be useful in an exploratory sense for stimulating the creation of hypotheses. A researcher with "a little knowledge" of statistics, who doesn't respect these

The Casting Aside of Human Agency. Human agency, or the usefulness of motive or ideas as explanatory concepts, or the assumption of individual rationality, has been explicitly or implicitly cast aside by some "scientistic" approaches in the social sciences since Durkheim (1981). This has been achieved in a number of ways: in sociology and anthropology by considering individual behaviour as deviations around a "norm" (see Hacking, 1990) understood in terms of non-human abstractions (structural and functional models based on mechanical and biological concepts); in political science by enclosing actual decision-making processes within a "black box", concentrating instead on inputs and outputs; in economics by abandoning the utility theory of its infancy in favour of "revealed preference" and technical market efficiency criteria like "pareto-optimality" instead of "utility maximization".

The early work of decision theorists like Lindblom, Simon and Etzioni, which showed that decision-makers often do not follow the specific decision process models that were thought to ensue from assumptions about human motivation and desire has been too quickly popularized to posit the "non-rational actor". This has been a setback for the social sciences; this vulgarization from a field which has shown itself to be one of the most dynamic of the social sciences (decision theory) enables researchers to convey the image of being contemporary, while continuing to treat the individual as an incomprehensible black box.

Conditions for Proof

The fundamental proposition of the social sciences in general is the rationality of the individual actor, based on her free will. To make an "irrational" decision is to choose a course of action for no reason, or without reason[8]. How can we conceive of "no reason" choices? As pushed by a body or

---

limitations, is truly dangerous! Example: simplistic correlations relating race with crime, intelligence quotients, and the dimensions of sex organs.

[8] This is a broad definition of rationality, which underlies most of the common definitions, many of which involve an assumption regarding the objects toward which reason is directed. The debate over rationality has often been confused by mistaking rationality —

bodies? If such a body is human, what propels that human?

If we decline to analyse the choices of humans under the proposition of rationality, we are left with a recursive chain of "causes" stretching indefinitely into the past. For, seeing life as "but a motion of limbs"[9] means a "human" science which explicates but does not explain; as will be discussed below, forecasts but does not predict; paralyzes but does not prescribe. In a science of "stimulus-response", according to the behaviourist Skinner, "The question of originality can be disposed of ..."[10].

What if such propelling bodies are human collectivities — social and economic institutions? Despite the problems with structuralism and functionalism mentioned here we should recognize the advance that such paradigms brought to our thinking: that human collectivities are a separate order of reality from individuals. Social institutions, once created by individuals, exert power over them, following a logic that cannot be understood as simply the sum of the desires of the individuals within those institutions. They are wholly dependent upon individual choices, however, and while they can be seen as having their own "needs" — equilibrium, adaptation, preservation, reproduction — their legitimacy and survival lie in the choices of individuals.

If such choices are irrational and hence incomprehensible, we are back to Cassirer's question: who made the machine and why? Are we left, then, with explanations like Austin's tautological explanation for sovereignty (Field, 1963), or Augustine's social institutions as given by God (1962)?

---

purposive, reasoned behaviour — with assumptions about the nature of purpose. For example, rationality has often been taken to imply Benthamite or Hobbesian selfishness and hedonism. This is, then, a "permissive" view of rationality in McClennen's terms; "it is essential to such a view of rationality that the agent's interests themselves — as expressed in terms of his preferences with regard to outcomes — are not subject to a critique in terms of principles of rationality" (1990, p. 84).

[9] Hobbes, cited in Bluhm, 1971, p. 288.

[10] Skinner, cited in Koestler, 1967, p. 14.

Empirically establishing the existence or non-existence of rationality is extremely difficult. To directly support a hypothesis of rationality, one must show that actions result from reasoned reflection. To support the non-existence of rationality, one must show that action does not result from reasoned reflection. The problem here is that irrationality is defined as the absence of rationality.

A hypothesis of irrationality can only be supported by the absence of such evidence, yet *the absence of evidence can only be taken as indicating irrationality if it were reasonable to expect to find evidence if the person were indeed rational.* E.g., lack of footprints in the snow indicates the non-presence of someone in the area only if there is snow. If instead the terrain is covered with running water, we cannot reasonably expect the absence of evidence to establish the non-existence of presence. Any evidence found is more likely to establish presence — e.g. a dropped watch — than the lack of it would support non-presence.

Irrefutability of Rationality Hypothesis

Do we ordinarily expect to find direct evidence of rationality, meaning the lack thereof supports the hypothesis of irrationality? No: rationality implies a non-observable thought process. We may find indicators of rationality, or support for it in experimental research (memos, reports of conversation, non-verbal cues such as chin-rubbing), but while the existence of observable indications will sometimes correlate with elements of the decision-making process (time taken, aspects considered, etc.), there is no reason to infer that the lack of indications implies lack of reasoning. This means that *to a large degree the hypothesis of rationality is non-refutable.*

According to Popper's criteria for refutability (1979) this should render it a poor hypothesis[11], but what is the alternative? The hypothesis of irrationality *is* refutable. This becomes, therefore, the "null" hypothesis. This doesn't

---

[11] However, in his "The Rationality Principle", Popper (1985, pp. 357-365) asserts that a strict rationality principle or law is refutable, on the basis of casual observation. I am not satisfied with this argument, nor with his solution to "water down" the principle.

mean that we *assume* or accept irrationality, but rather that we, as researchers, search for evidence that humans are *rational* since the likelihood of finding evidence of irrationality is virtually impossible[12].

Where does this leave attributions of irrationality? Generally speaking, we cannot say people are irrational: in many cases we will find evidence of rationality; in others that we have not seen evidence of rationality, implying nothing, since we didn't expect to find such evidence or haven't yet looked for it. In very few cases, the logical possibility remains that we will expect to find evidence of rationality but did not, thereby justifying resting with the null hypothesis of irrationality, while not proving it.

Testing for Rationality

I do not want to go too deeply into the practical problems of gathering evidence regarding thought processes, but there is one aspect regarding the theoretical formulation of any such experiment that must be mentioned: evidence of extremely little thought given to a *particular* decision does not establish a *prima facie* case for irrationality. Decision theory based on observation of managers' decision-making processes questioned the simple maximization/optimization models we had previously posited (e.g., Etzioni's "muddling through").

Rather than supporting the hypothesis of irrationality, such processes may be reasonable ways to reduce the time and effort to make decisions. When placed within an economic framework, it no longer appears "rational" to spend an inordinate amount of time making decisions, especially routine decisions (Elster, 1990).

Meta-Decisions. Another way to view this issue is to place particular decisions within a **meta-decision** framework. Rather than seeing an individual as going from one discrete decision to another, one needs to place a decision in a large enough temporal space such that the multiple objectives of the individual

---

[12] According to William James: "*The true opposites of belief,* psychologically considered, *are doubt and inquiry, not disbelief*" (1970, p. 158).

have a chance of revealing the complexity of any particular decision. Reducing to an acceptable level time spent making a decision might be one such objective; other objectives can be seen as "conflicting". It is a basic principle of decision theory that unless objectives are mutually independent, which they rarely are (e.g., given substitutability, or time and money constraints) they cannot be simultaneously optimized[13].

Hence, decisions which seem perverse might only be understandable with respect to objectives which do not obviously bear on the problem, or with respect to a larger strategy. In order to save decision time the individual may have made a meta-decision as to how decisions would be made in certain classes of problems.

Evidence that a decision was "wrong", even given the individual's objectives, does not necessarily imply an irrational decision process: any such process will implicitly carry with it some expected proportion of wrong decisions. A decision process that maximizes the number of "right" decisions is probably inefficient, as it implies a much greater amount of effort in making decisions than under the short-cuts adopted by the rational person. Indeed, it is theoretically possible, if unlikely, that *obvious evidence of rationality might be inversely correlated with actual rationality.*

Human Complexity. The search for rationality therefore implies a radically different perspective from the "mechanical" sciences: the gaze is directed outward to the context of the individual and upward through hierarchies of belief and value, rather than to the causal links and chains. The latter implies a linear process of description, even if "multivariate", rather than *explanation* of complex behaviour via analytical approaches having the flexibility and subtlety adequate to capture and convey the rich, complex, strategic nature of human behaviour[14]. One will not find meaning in a Hobbesian "motion of limbs" for, as Bateson (1972, p. 316) notes, "the mental characteristics of the system are immanent, not in some part, but in the system as a whole."

---

13. See Miller and Starr, 1960, pp. 39-52.

14. See Morin, 1990; Wallace, 1974.

Toward a Teleological Human Science

Explanation: Conditions and Purpose

If rationality means having reasons for the choices one makes, rationality can only be understood with respect to *purpose*. If we attempt to view the reasoning process as a matter of creating a narrative around causal forces, we might ask why such an ability to construct narrative would arise if it only provides a story to describe actions which we are propelled to do.

The answer is obviously to be able to choose among alternative courses of action given those causal forces. If we have the free will to do so, than "causal forces" have to be viewed instead as "conditions", and the choice must be seen as having some independence from causal forces, meaning the choice is to some extent defined not by conditions but by purpose[15].

Explanation in the human sciences, therefore, must centre around purpose: decisions need to be understood as a choice among alternatives. To refer to a variable as "causal" is, in the human sciences, to admit that all we have been able to establish is a correlation[16]. When the role of a variable in shaping conditions, and therefore altering the nature of alternatives, is understood, then its role as a factor in the decision-making process is properly situated.

It is never useful to employ, unqualified, the concept of causation in the human sciences; even saying that an agent's choice "caused" an action begs explanation. Since the "forces" of human action are purposeful choices, the

---

[15] Charlton (1986) said that consciousness of causality and teleology is "consciousness of two aspects of ourselves", but "what we are conscious of in understanding causally is not the same as what we are conscious of in understanding teleologically" (p. 101).

[16] C.S. Pearce, "The Doctrine of Necessity Examined", cited in Hacking, 1990. The establishment of temporal precedence is a necessary but far from sufficient condition for establishing "causality", leave alone explanation.

human sciences must be a study of decision-making in a context of constraints, opportunities, uncertainty, and values[17].

Prediction as Prophecy

Alasdair MacIntyre (1973, 1977) discussed challenges with respect to the object of social science — or knowing in general — that if not unknown in the natural sciences are much less problematic. Whether the seeker of understanding is Hamlet facing a personal epistemological crisis, or the social scientist trying to comprehend human behaviour, the problems raised, according to MacIntyre, stem from the unpredictability of human behaviour, based as it is upon reflection, meaning and choice (greatly complicated by the strategic interactions between agents). "Ambiguity, the possibility of alternative interpretations, becomes a central feature of human character and activity" (1977, p. 459).

One alternative to prediction, prophecy, was mentioned by Nisbet (1969): whereas a prediction is essentially deterministic in nature and implicitly assumes the benigness of its own impact, a prophecy counts on a long-range vision of what could be, perhaps of what should be, of human agency, and the role of the prophecy itself in helping to fulfill itself.

Social vs. Natural Sciences. MacIntyre's concerns, and the concept of prophecy, capture some of the essential differences between prediction in the *social* as compared to the *natural* sciences:

- The "conditions" out of which the future enfolds are ordinarily far less identifiable or controllable.
- The making of a prediction can substantially alter human behaviour, rendering that predicted more or less likely. Unless one can isolate the

---

[17] It is important to distinguish between a "science of choice" as Robert Mundell characterized economics (cited in Buchanan, 1987, p. 35) as purely the "methodological" study of decision-making *processes* and as a science that actually incorporates process models in the study of human value and purpose; the failure of economics to do the latter leads Buchanan to question whether economics is really a science of choice.

impact of a prediction — extremely problematic as even obscure predictions can have a snowball effect[18] and because of the difficulty of separating the impact of predictions from other factors considered in decision-making — it is very questionable as to whether one can place any confidence in a prediction "track record" as an indication of how "good" the predictions were[19].

- Decision criteria can so substantially change as to be not only unpredictable, but even incomprehensible, from one time to another. This can be seen as a problem of emergence (Polanyi, 1967): the evolution from one system to a more complex system cannot be predicted on the basis of knowledge of the less complex system (see Dawkins, 1989). As consciousness reorganizes itself, what is "real", and what is "desirable", can change radically. Take, for example, the purpose of science itself: could linear causal explanations predict the future of science if we move from a utilitarian approach in research to the "age of the new scientific spirit" described by Bachelard (1969), whereby solutions to problems are sought in order to ask better questions?

What is the purpose of prediction in the sciences? One of the roles of prediction is to verify hypotheses generated from past observations. As discussed above, this is problematic in the social sciences because of the relative inability of the scientist to avoid "contaminating" the observed. Aside from this problem, the ability of any social science to predict non-trivial outcomes is likely to be always so limited that it remains doubtful that simple prediction should be seen as a primary objective of the social sciences.

---

18. This is particularly so if systems of human behaviour are seen as "chaotic", with extreme sensitivity to initial conditions, as described by Gleick (1987).

19. A partial exception to this can be found in controlled experiments. In this case, however, successful prediction might be trivial, if the experimental results have little external validity (i.e., application to human behaviour in the "real world"). See Cohen and Nagel, 1934, pp. 208-212.

A Teleological Human Science as the "Art of the Possible"

What distinguishes scientific from subjective knowledge is its justifiability. The grounds for justifiability are a matter of community norms (Polanyi, 1967). Perhaps knowledge in the natural sciences may still be justified with respect to its closeness to the "real", but it is being increasingly recognized that the human sciences can seek no such justification.

I propose a fundamentally different paradigm for the social sciences: that being the *art of the possible* rather than the science of the real. A possibilist science has as its aim *prescription* rather than prediction: by referring to what is "right" or "good", and by trying to delineate the options open to us, this science would aim to provide guidance as to which roads to embark upon and how to correct our direction as we progress.

This paper has had to restrict itself to the most abstract level in discussing the direction that the social sciences should take; unfortunately I have not been able to discuss except at the most general level the methodological implications of a teleological science, and whether application of such a science is practicable. I will, however, provide one example from the field of mental health where a teleological approach might be seen as useful.

I have analyzed determinist explanations with respect to deinstitutionalization (McCubbin, 1994); the recent skepticism in the literature which sees the cause of the deinstitutionalization process, and the failure to establish a new biopsychosocial community care model, as inevitable results of economic forces, in itself seems to be contributing to a feeling of disillusionment which has followed various attempts at social reform begun during the 1960s. This disillusionment has the effect of a *self-fulfilling prophecy*.

Although our understanding of the past may have been deepened as a result of the determinist arguments, I am not convinced that they have contributed as much to the future as a teleological approach might. If the "cause" of failure is identified as something easily surpassed, then the deterministic approach in that case would be useful.

However, *the very nature of deterministic argument results in emphasis on the inevitability of the result*: A causal explanation is not usually seen as forceful unless the "forces" were shown to be extremely powerful. These arguments tend not to show how history was made as a series of choices, which would be a more useful approach in guiding our decision-making for the *future.*

A teleological approach might, for example, evaluate which barriers lie in the way of certain valued objectives — e.g., regarding reinsertion of distressed or offending persons within the community. The "lessons of history" are brought in not to provide "laws" with respect to the future, but to better help us understand our environment and assess our alternatives.

One might conclude, on the basis of history, that the power of psychiatry is so strong, and that the present interests of the institution of psychiatry are so anathemic to the objectives of a biopsychosocial model, that a major shift in the mental health system is extremely unlikely. Policy changes reducing the legal power of psychiatry appear to have made little difference. Acceptance of this hypothesis implies a static policy, or further attempts to limit the power of psychiatry.

But how much of the acceptance by various parties — family members, academics — of the biochemical theory of mainstream psychiatry is based on strategic choice that has *little to do with the credibility of psychiatry* (I think, for example, of the role of pharmaceutical companies as funders of research, and the desire of family members to avoid blame as dysfunctional families)? In such a situation a decline in the real power of psychiatry might be hard to detect.

The Role of Values in a Teleological Science

Would we not have a better chance of identifying *future* rather than past constraints by evaluating the actual choice problems faced by the various parties, past and present, with respect to constraints as they would see them, and with respect to their own values? Would we not be more useful, as social scientists, by openly and actively participating in society's formulation of values and objectives? How can we hope to make rational choices as a society unless we constantly renew and maintain in the forefront our own values?

A teleological science is not "value-neutral": a teleological scientist will be at least as ruthless in identifying and pursuing "true" values and objectives as she is with the "facts". If there is no place for relativism with regard to the data of the natural sciences, there is no place for value relativism in the social sciences: for not only would a "value-free" human science have no *meaning*, it would have no *use* in understanding, predicting or prescribing human behaviour.

Our free will necessitates rationality, the existence of purposeful behaviour implies rationality and free will, and purposeful behaviour can only be understood in the context of value. By the very act of "describing" behaviour we substantially influence it; the values underlying what a scientist chooses to describe, and how to describe it, convey, explicitly or not, a statement about the values she holds, and therefore those that she will propose[20].

Conclusion: A Teleological Science As A Uniquely *Human* Science

By focusing on:

- possibility rather than necessity,
- decision-making rather than determinism,
- final cause rather than initial cause,
- teleology rather than ontology,
- beliefs rather than facts, and
- values rather than forces,

the human sciences might aspire to the maturity with respect to their object of study as that which is commonly accorded to the natural sciences with respect

---

[20] In this paper I have attempted to delineate the connections between free will, rationality, and human behaviour as a teleological phenomenon. In doing so it has become clear how intimately involved the social scientist is with those he "observes". The next step in the analysis would be an evaluation of the concepts of consciousness, mind, and collective consciousness. It may be that given the high degree of social interconnectedness, and the importance of shared fundamental values, that rather than seeing individuals as "atoms in a context" , a perspective of society as having collective consciousness would be more fecund for certain problems, aside from the ontological justification for such a view.

to their objects of study. While social science can be enriched by judicial use of the concepts developed in the natural sciences (e.g., cybernetics, chaos theory) — perhaps as metaphors or abstract process types — there is no reason to assume that models useful for explaining the phenomena of plants and rocks can explain human behaviour, and even less, guide our future choices.

A teleological science would be a uniquely human science based on the fundamental element of what it is to be a human being: activity purposefully following our values formed and reformed by our own developing consciousness.

References

Augustine, Saint. (1962). *The political writings of St. Augustine* (Ed., with intro., by H. Paolucci). Chicago: Gateway.

Bachelard, G. (1969). *La formation de l'esprit scientifique: Contribution à une psychanalyse de la connaissance objective.* Paris: J. Vrin.

Bateson, G. (1972). *Steps to an ecology of mind: The new information sciences can lead to a new understanding of man.* New York: Ballantine.

Bertalanffy, L. von (1981). *A systems view of man.* Boulder: Westview.

Bluhm, W. T. (1971). *Theories of the political system.* Toronto: Prentice-Hall.

Buchanan, J. M. (1987). *Economics: Between predictive science and moral philosophy.* College Station: Texas A&M University Press.

Campbell, J. (1982). *Grammatical man: Information, entropy, language, and life.* New York: Simon & Schuster.

Cassirer, E. (1944). *An essay on man.* New Haven: Yale University Press.

Castoriadis, C. (1978). *Les carrefours du labyrinthe.* Paris: Seuil.

Charlton, W. (1986). Knowing what we think. In L. Stevenson, R. Squires, and J. Haldane (Eds.), *Mind, causation & action* (pp. 86-102). Oxford: Basil Blackwell.

Cohen, M. R., and Nagel, E. (1934). *An introduction to logic and scientific method.* New York: Harcourt, Brace & World.

Dawkins R. (1989). *The selfish gene.* Oxford: Oxford University Press.

Durkheim, E. (1981). *Les règles de la méthode sociologique.* Paris: Presses Universitaires de France.

Elster, J. (1990). When rationality fails. In K. S. Cook and M. Levi (Eds.), *The limits of rationality* (pp. 19-51). Chicago: University of Chicago Press.

Etzioni, A. (1988). *The moral dimension: Toward a new economics.* New York: The Free Press.

Field, G. C. (1963). *Political theory.* London, U.K.: Methuen.

Foucault, M. (1954). *Maladie mentale et psychologie.* Paris: Presses Universitaires de France.

Gardenfors, P. (1992). The dynamics of belief systems: Foundations versus coherence theories. In C. Bicchieri and M. L. D. Chiara (Eds.), *Knowledge, belief, and strategic interaction* (pp. 377-396). Cambridge: Cambridge University Press.

Gleick, J. (1987). *Chaos: Making a new science.* New York: Penguin.

Hacking, I. (1990). *The taming of chance.* New York: Cambridge University Press.

Hobbes, T. (1983[1651]). *Leviathan.* Glasgow, United Kingdom: William Collins.

James, W. (1970). Belief and Perception of Reality. In H. S. Thayer (Ed.), *Pragmatism: The classic writings* (pp. 157-174). New York: Mentor.

Koestler, A. (1967). *The ghost in the machine.* Toronto: Macmillan.

Koons, R. C. (1992). *Paradoxes of belief and strategic rationality.* New York: Cambridge University Press.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

MacIntyre, A. (1973). Ideology, social science, and revolution. *Comparative Politics, 5,* 321-342.

MacIntyre, A. (1977). Epistemological crises, dramatic narrative and the philosophy of science. *The Monist, 60,* 453-472.

McCubbin, M. (1994). Deinstitutionalization: The illusion of disillusion. *Journal of Mind and Behavior, 15,* 35-53.

McClennen, E. F. (1990). *Rationality and dynamic choice: Foundational explorations.* New York: Cambridge University Press.

Miller, D. W., and Starr, M. K. (1960). *Executive decisions and operations research.* Eaglewood Cliffs, New Jersey: Prentice-Hall.

Morin, E. (1990). *Introduction à la pensée complexe.* Paris: ESF Éditeur.

Munn, A. (1960). *Free will and determinism.* Toronto: University of Toronto Press.

Nisbet, R. A. (1969). *Social change and history: Aspects of the western theory of development.* New York: Oxford University Press.

Pearce, W. B. (1989). *Communication and the human condition.* Carbondale: Southern Illinois University Press.

Polanyi, M. (1967). *The tacit dimension.* Garden City, New York: Anchor.

Popper, K. R. (1979). *Objective knowledge: An evolutionary approach.* Oxford: University Press.

Popper, K. R. (1985). The rationality principle. In D. Miller (Ed.), *Popper selections* (pp. 357-365). Princeton, New Jersey: Princeton University Press.

Sciama, D.W. (1959). *The unity of the universe.* Garden City, New York: Anchor.

Teilhard de Chardin, P. (1955). *Le phénomène humain.* Paris: Seuil.

Wallace, W. A. (1974). *Causality and scientific explanation*. Ann Arbor: University of Michigan Press.

Wright, G. H. von (1974). *Causality and determinism*. New York: Columbia University Press.

General System Theory as a Postmodern Epistemology for the Social Sciences

Michael McCubbin

Groupe de recherche sur les aspects sociaux de la santé et de la prévention
(GRASP), Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal
(Québec), Canada, H3C 3J7

[redacted]

Keywords—General System Theory, social science, philosophy, epistemology, methodology, paradigm, system dynamics and process, chaos, postmodernism, critical theory, teleology, collective consciousness, emergence

Acknowledgments—My thanks to Drs. D. Cohen, F. Lesemann, and N. Almeida-Filhõ for their comments, and to GRASP and the Université de Montréal for financial and other support while this paper was being written.

29 January 1998

General System Theory as a Postmodern Epistemology for the Social Sciences

When I went out this morning,
I reflected on man's humanity
In regards to the entire solar system
Which may or may not explore
On it's own
The entire length of the Hollow
Known as the universe
Sometimes I feel like the only part of it I am is a piece
Of dust of the source-of-All
Which is part of the Self-Actualizer
Be that as it may, I will still not
Cease to pray
That I has a spot
To say nothing about Today[1].

This paper discusses the pertinence of General System Theory (GST) as an epistemology for the social sciences which responds to the postmodern critique of modernistic epistemology. Part I places GST within the context of the modern/postmodern narrative, provides a history of GST development, summarizes its key concepts, and indicates some of the cutting edges of the field. Part II, after a brief indication of the relationship of GST to the natural sciences, reviews some of the thrusts GST has made within a variety of social sciences. Part III concludes the paper, with a discussion of the promises and limitations of the GST approach for social sciences, particularly in terms of understanding social problems.

---

[1] This poem (Eckler, 1994) was provided without interpretation or context in a newsletter published by a mental health patients' council. It addresses some of the more philosophical and personal implications raised by GST; the reader may wish to return to it after reading this paper. I have not corrected the apparent mistakes.

When seen primarily as an epistemology, GST can withstand common critiques, which are actually aimed at various empirical applications. As an epistemology GST is seen as able to provide a synthetic framework for conceptualizing the research process. When combined with empirical theory and assumptions, it can be a fecund source of hypotheses, particularly with respect to dynamic teleological behaviour. Systems thinking has interesting implications for explanation and understanding in the social sciences as dynamic or developmental processes.

Part I: An Overview Of GST

A Brief Definition of Systems Theory

Ludwig von Bertalanffy, in articles he originally published almost 50 years ago, defined a system as "a complex of interacting elements" (1968b, p. 55). This conception has been widely accepted by those explicitly applying general systems theory (GST)[2]. Von Bertalanffy saw GST as pure theory, a science which would proceed via the elaboration of axioms given definitions and assumptions; i.e., given certain assumptions, the behaviour of an *abstract* system can be *deduced.* Application of GST to "real-world" problems therefore requires the identification of isomorphisms between theoretical and empirical systems in order to develop hypotheses or predictions.

The basic principle of GST, and its most significant contribution to postmodern science, is the recognition that a system of interacting "parts" cannot

---

[2] Although Acker defines system as a "set of interrelated elements each of which is related directly or indirectly to every other element, and no subset of which is unrelated to any other subset" (cited in Shakun, 1981, p. 235), which seems implied by von Bertalanffy's definition, and Mesarovic and Takahara (1975) define system in set-theoretic terms as a set of relations among a defined set of objects. This latter definition has advantages in that it emphasizes "system" as a matter of *observer* denotation, and in that it makes no *a priori* assumptions about the existence or form of relations, thereby providing the most general definition, which can then be elaborated for specification of system *types*.

be understood by separate *analysis* of the parts because of their complex interactions and because "purpose" or "meaning" may only be immanent in the whole[3]. Hence "systems thinking" is *synthetic*, dialectical and emphasizing deduction, rather than analytic, emphasizing induction (Sutherland, 1973). The latter forms of analysis are only appropriate for static closed (or almost closed) systems characterized by simple structures, lacking complex interaction among the elements.

GST is not seen here, nor by its mainstream theorists, as intrinsically a theory about reality; rather, it is an epistemology that structures how we think, not only about reality, but also about thinking itself; in effect, in its abstract form GST is a theory of knowledge. A system model of reality is created by progressively adding assumptions about the "system" we wish to explain. A number of such assumptions have generic types, in that they have been discussed or defined for species of abstract systems by GST theorists (structures, processes, degree of openness, complexity). One of the most interesting aspects of GST is that it enables cross-system or cross-level comparison on the basis of systemic similarities and differences. This implies a major advance in methodology if all "meaning" is seen as attribution of similarities and differences among abstract or concrete objects, and if we can avoid overly petrified ways of seeing "systemness".

While probably any set of explanatory factors can be brought in and modeled without GST, the latter offers to facilitate that process as one that is systematic, thorough, and self-conscious. Von Bertalanffy (1968b) referred to GST as a meta-theory akin to mathematics. GST provides a way to abstract from reality, potentially simplifying it while at the same time capturing its multi-dimensionality and recognizing its constructed nature, but in terms precise enough to enable deduction and subsequent empirical testing[4].

---

[3] Nor can the "parts" — or subsystems — be understood or anticipated (except in special cases) by reference only to the "whole" (Morin, 1990); for example, von Bertalanffy (1968b) describes the concept of equifinality whereby different systems achieve similar results (or where a perturbation does not alter the end result).

[4] I will provide a heuristic explanation for the advantage a simplifying theory may have over pure observation. The formula for the volume of a cylinder is purely mathematical, part of a

Systems thinking is liable to lead to the empirical observation that entirely closed systems can exist only in abstraction. Even test-tube chemical reactions are slightly contaminated by environmental influences (although perhaps not enough to justify rejecting an assumption of closure). The recognition that systems are open, particularly human biological and social systems, helped lay the groundwork for a revolutionary change in the way the sciences approach the "object" — or, more properly speaking, *phenomena* — of interest. The consequences of "open-ness" are profound, as will be discussed below.

Systems Theory in the Context of Postmodernism

At the risk of repeating what has become a cliché, the paradigm of "modernistic" science is decaying. Systems thinking has been an early and very important component of the critique of modernistic science and, perhaps more importantly, to the development of new paradigms and theoretical approaches.

The failures of modernistic epistemology and methodology, along with their implicit values, have been so frequently recounted that reading and repeating them becomes tedious. Nevertheless, most papers explicitly using GST and published in the traditional disciplinary journals, as well as many others from various postmodern streams, repeat the litany of modernistic inadequacies, for the simple reason that despite lip-service being paid to "complexity", "context", and "subjectivity", the bulk of research in most disciplines tenaciously adheres to modernistic perspectives and methods; indeed, it appears that some social science fields have been enhancing their modernistic orientation[5] in order to achieve the

---

"tautological" system. While one might obtain a reasonable volume estimate by filling the cylinder with liquid (after sealing the bottom, of course), calculating the change in the "closed" room's humidity after the water evaporates, and applying the additional formulas necessary to arrive at an estimate of the liquid's volume, it is much easier to measure the height and radius of the cylinder and apply the formula for volume of a cylinder. While the first method contains no assumptions about the cylinder's perfectness as a cylinder, it is less direct than the second method, and inevitably also contains assumptions, explicit or implicit (e.g., about the "closedness" of the room).

[5] The argument that the human sciences have become more rather than less modernistic receives better support over a historical perspective of the last half century, in which we saw the

scientific cachet which, while becoming tarnished, continues to shine brightly — especially within society in general.

The following discusses some of the problems most frequently identified with modernistic science, to which postmodern approaches, including GST, aim their critique.

Linear Causality. Chains of linked bivariate causal relations are no longer considered sufficient for explanation and prediction. Insofar as events are "determined", they have *multiple* causes. Much of postmodernism rejects or modifies notions of causality. In general, systems philosophy finds the concept an impediment to understanding open systems, particularly self-organizing dynamic systems. *Final* cause or teleological conceptions are frequently brought in: von Bertalanffy's "equifinality" (1968b), or Shakun's (1981) "evolutionary systems design" based on changing value goals within a collective consciousness.

Separation of Object and Subject. Foucault (1966) and Bachelard (1969) have convincingly posited as necessary for the development of modern science the ability of the observer to see the observed as "separate". The perception of separateness was posited as a necessary stage of the *advancement* of science seen as the communication of symbols within a common language about a common "reality"[6] that enabled the building upon — or, in postmodern terms, the

---

rise of behaviourism in psychology and the widespread use of quantitative methods in most social science disciplines; it is not clear to what degree the increased use of "interpretive" methodologies over the past decade or so is a harbinger of long-term change, rather than a mere oscillation.

Hacking (1990) suggests that a new stage has been reached this century by the use of probability theory, which today underlies much quantitative research in the social sciences. McCubbin (1994a) suggests that this "taming of chance" does not fully respond to the critique of modernism, suggesting that we need to progress beyond the science of probability to the science of *possibility*, incorporating conceptions of human agency and teleology.

[6] I will use the term "reality" many times in this paper without quotation marks; however quotation marks should be assumed, indicating that "reality" which we can assume, understand, or perceive. Whether anything is in fact real is another question, on which I am agnostic. One problem is the seeming impossibility of determining or conceiving of an *objective* criterion for

complexification of — symbolic descriptions.

However, as with any paradigmatic precept, the conception of separateness was useful for some problems but not others; the recognition by GST that most systems are open, especially human systems, renders inevitable an epistemology that recognizes the influence, "embeddedness" (Granovetter, 1985), and enmeshment of the observer with the observed (Checkland, 1981). Dobuzinskis (1992), for example, presents systems-theoretic models of the policy process which incorporate the influence of the policy analyst.

Positivism. This critique questions the (attributed) view of modernistic science that it can gradually progress *towards* apprehension of reality by applying increasingly refined techniques and tools of observation and measurement. The growing self-consciousness of modernistic science, as manifested in its epistemological ideology, contained the seeds of its own destruction — or, in systems terms, *emergence* (Polanyi, 1967) to a new level — as even early philosophers of scientific method like Kant (1958) recognized the distinction between what is perceived and what "is".

The epistemological challenge received new impetus from the development of the cognitive sciences, and its important contribution from Gestalt theory (Köhler, 1947). Gestalt theory posited that what we see with our own eyes is mostly based on imagination, expectation, and prior experience. The response to the critique of positivism varies among systems schools as in postmodern streams generally. Mainstream GST emphasizes the abstract nature of the models which are used to attempt to describe reality (if indeed "reality" is acknowledged as a meaningful conception). By 1950 von Bertalanffy had warned (1968a) against the misuse of systems thinking by what later came to be lamented as the "reification" of abstract systems (e.g., critiques of Parsonian functionalism, as described by Bailey, 1993, 1994)[7].

---

separating that which is truly real from that which is not.

[7] "... [W]e are careful not to slide into 'biologism'.... The organismic conception does not mean a unilateral dominance of biological conceptions" (von Bertalanffy, 1968b, p. 88; written prior to 1950).

Object of Study. Explanation based on physical things propelled by forces has become less satisfactory. Ironically, physics was one of the first sciences to realize that the limitations of physicalist science had begun to loom large; in many respects the "empirical" social sciences are not emulating the natural sciences of today but of a century ago. Von Bertalanffy (1968b) described Newtonian science as reaching its limits when it began to direct its attention to the microscopic and macroscopic or cosmic levels. He not only stressed the importance of shifting attention from elements in structure to elements in *interaction*, but also that the human systems we are dealing with are increasingly *symbolic* (see also van Gigch, 1990).

Contextuality. As science directed its attention to new or redefined problems and objects of study, including humanity itself, and sought greater explanatory and predictive power, the construction of problems became more complex, to the point that scientific method needed to change. A common response to this critique has been to create "multivariate" models with multiple "determinants". This is often done without explicit prior theory — the computer is allowed to create a model based on correlation (*linear* regression). The resulting "models" statistically describe *samples*, and, we *hope*, the populations from which they are drawn, but provide little in the way of understanding or external validity. Correction for "interaction effects" can result in trivial models which provide a "good fit" but suffer from an analytical viewpoint[8].

"Thick thinking" (Bibeau, 1988) needs not only to ensure that the multiple dimensions, layers, and folds are represented by the inclusion of variables, but that forms of interaction are *modeled*, since in the "real" world relationships are almost never linear (or log-linear), and that while interactions may become quite complex, they may sometimes be simply characterized given an understanding or theory of system dynamics (e.g., fractals characterizing chaotic systems; see Gleick, 1987). A

---

[8] See Kerlinger (1986) for a discussion of the importance of hypotheses in conducting empirical research; e.g.:

> ...[I]t is quite conceivable that the relation is fortuitous or even spurious. If, however, he had hypothesized the relation on the basis of theory, the investigator could have greater confidence in the results. Investigators who do not hypothesize relations in advance, in short, do not give give [sic] the facts a chance to prove or disprove anything (p. 20).

number of promising, though not well known, non-linear statistical methodologies have been developed by systems methodologists (Hofferbert and Schäfer, 1982).

The recognition of environmental complexity means that the researcher can be more explicit about the limitations of a particular model and less likely to make egregious errors; it does not imply that operationalized models or theories must themselves be complex! (How often have we seen research which attempts to deal with "complexity" by measuring dozens, or even hundreds, of variables?) The crucial aspect of GST is epistemological, not methodological; given a good *explicit* epistemology, the researcher is more likely to pick an *appropriate* methodology, practical in the circumstances. The generalizability of the results can then be more clearly understood.

As we (apparently) move into the postmodern era, various theories and approaches have been competing to provide critiques and narratives, and in most cases alternatives, to modernistic thinking. In its systematic and coherent treatment of the problems of complexity and context, systems thinking addresses many of the issues raised by the postmodern critique, and has influenced other postmodern approaches.

A Brief History of the Development of Systems Theory

Ludwig von Bertalanffy is clearly the "father" of GST, although a number of concurrent streams of research and thought contributed to the development of systems thinking and specialties. Von Bertalanffy (1968b) identified a number of writers in the 1920s as laying some of the early groundwork leading to GST, including Köhler (*gestalten*), Lotka (communities as systems), Whitehead (organic mechanism), Cannon (homeostasis), and Claude Bernard (organismic conception).

Following von Bertalanffy's work in developing an organismic conception in biology (in the context of a debate between mechanicists and vitalists), he developed theories of "open systems" and "steady states", later generalized into what he called "General Systems Theory", which he verbally presented in 1937. He published his ideas regarding open and closed systems, "equifinality", and steady state in 1940 in an article describing the organism as an open system (reprinted in von Bertalanffy, 1968b, pp. 120-138). By 1950 he had published mathematical generalizations of the basic concepts of GST (1968b, pp. 54-88). He published the seminal article for the field in 1962 (most frequently cited as 1968a).

Two of the most well-known applications within the systems field, today often mistakenly taken to be the systems approach in general, are cybernetics and structural/functionalism. These streams appear to have emerged largely independently from the nascent GST literature.

Cybernetics. Norbert Wiener introduced the term "cybernetics" in 1947, to describe a field he and colleagues had been working in, having "become aware of the essential unity of the set of problems centering about communication, control, and statistical mechanics, whether in the machine or in living tissue" (1948, pp. 11-12). Wiener elaborated a number of systems concepts including feedback, oscillation, and self-organization, and dealt with non-linear, multivariate problems and dynamic systems.

Von Bertalanffy (1968b) noted a number of characteristics of cybernetics which render its models as special cases in systems theory, most of which can be related to the closed system nature of cybernetic models. Hence cybernetics can explain homeostasis (self-reproduction and corrective adjustments) but not "heterostasis" (steady state dynamic change in the nature of the system, including increasing complexity [negentropy] and goal changes)[9]. Cybernetics remains very productive, particularly in its use of information theory, in explication of goal-seeking behaviour, and in modeling complex feedback mechanisms. This field provides competition and enrichment to GST philosophers, with often more mathematical or applied contributions[10]. The application of cybernetics to

---

[9] Gray and Rizzo (1969) attribute "heterostasis" to Menninger:

> Menninger goes beyond the principle of homeostasis and feels that the phenomena of growth and development demand an entirely different principle, namely, heterostasis, the progressive moving away from the status quo (p. 19)

It is unfortunate that the term "heterostasis" is rarely used, since the term "dynamic", often used instead, could apply to any active homeostatic system, including those that maintain static boundaries and level of complexity.

[10] See Wiener (1950), for an introduction to cybernetics and its potential contribution to humanistic science.

artificial intelligence provides a thought-provoking approach to resolving the paradox of trying to understand humanity by humanity observing itself (see Foucault, 1962; Morin, 1990).

Structural/Functionalism. The structural/functional tradition in sociology contained many elements found in GST. Roots for this tradition can be found in Marx, Hegel, Durkheim (1981), and Weber. It reached its apotheosis in the work of Talcott Parsons (1951). Parsons provided a systematic analysis of abstract social structures, advancing thought on social system/subsystem functions, norms, rules, roles, equilibrium, and reproduction. Although in sociology structural/functionalism has been virtually equated with systems theory, it should be considered as based on special cases of system models (relatively closed and homeostatic) in the particular application of specific theorists.

GST itself is devoid of empirical content — it is a *metadiscipline* which makes available abstract models for application by disciplinary researchers. In my view the field of sociology has suffered from its apparently wholesale rejection of Parsonian approaches; these approaches were systematic and fecund, and could be extended in response to criticism (overly holistic, structure-oriented, static, closed). The integration of Parsonian inspired models with certain newer topics in sociology such as reciprocity theory (see Sahlins, 1965), strategic approaches (Crozier and Friedberg, 1977), and institutional theory (Meyer and Scott, 1992) would go a long way toward satisfying past critiques — and result in a synthesis benefiting from the maturation of GST since the Parsonian era.

GST as a Metadiscipline. Some other fields have contributed to the elaboration of systems theory: decision and game theory (Rapoport and Chammah, 1965), information theory (Campbell, 1982), and chaos theory (Gleick, 1987). As a metadiscipline, GST is engaged in continuous cross-fertilization with fields and disciplines ranging from biology, mathematics, statistics, and physics to political science, psychology and psychiatry, nursing and sociology.

The number of researchers explicitly identifying with GST appears to have stabilized in the 1970s after twenty years of rapid growth. Since that time the number of published papers overtly applying GST in the disciplinary journals may have declined, at least in proportion to the explosion of published articles, being displaced by more fashionable postmodern approaches and by some resurgence of strictly empirical work[11]. However, in some disciplines systems thinking has become so entrenched in the disciplinary paradigms that overt identification with GST has been dispensed with, and in other disciplines systems thinking continues its diffusion under the guise of disciplinary applications.

The early promoters of GST saw it as an inherently interdisciplinary or metadisciplinary field which could be drawn on by the various disciplines. While it is inevitable and in some ways desirable that each discipline develop specialization and its own language, it is important that they develop bridges to the metadisciplines. Otherwise, important systems-influenced developments within the disciplines run the risk of ossifying.

This danger is also incipient in the metadisciplines such as philosophy, mathematics and GST: they need inspiration from the disciplines with their more practical and empirical concerns. The metadisciplines need to be able to communicate their advances to the disciplines. The systems field now contains many specialized interdisciplines (loosely under the umbrella of systems and operations research), and some theoretical approaches which show signs of over-protectiveness (e.g., Living Systems Theory, very frequently published in the journal *Behavioral Science*). To some degree system sub-fields have lost the ability to communicate with each other and with the disciplines[12].

---

11. These comments may not be applicable beyond North America; even in the English language systems journals many contributors are from east and west Europe, China and Japan.

12. Ackoff (1974) distinguishes between the synthetic thinking of a metadiscipline and the analytical thinking of a discipline. He expresses regret that many interdisciplines lost the interdisciplinary concept:

> The interdisciplines sought recognition and status by emulating the disciplines and professions... [and] began to identify themselves with the instruments which they developed and used — that is structurally — rather than with what these instruments were used for — that is, functionally. They began to ... contemplate their own navels... (p. 18).

Paramount Principles of General Systems Theory

There is a fair amount of agreement in the literature of GST and its explicit applications about the most basic concepts in the field; disagreement arises in attribution of system type to empirical situations. The basic definitions, concepts and principles were outlined by von Bertalanffy (1968b). Ackoff (1974, p. 3) provides a succinct definition of a system, which includes properties which apply by definition or axiomatically:

> A system is a set of two or more elements of *any* kind; for example, concepts ..., ideas ..., objects ..., or people .... Therefore, it is *not* an ultimate indivisible element but a whole that can be divided into parts. The elements of the set and the set of elements have the following three properties:
> 1) The properties or behavior of each element in the set has an effect on the properties or behavior of the set as a whole. ...
> 2) The properties and behavior of each part and the way they affect the whole depends on the properties and behavior of at least one other element in the set. Therefore, no part has an independent effect on the whole and each is affected by at least one other part. ...
> 3) Every possible subgroup of elements in the set has the first two properties... Therefore, the elements cannot be organized into independent subgroups. A system cannot be divided into independent subsystems. ....
>
> Because of these properties a set of elements that form a system always has some characteristics, or can display some behavior, that none of its elements or subsystems can .... A system, viewed structurally, is a divisible whole; but viewed functionally it is an *indivisible whole* in the sense that some of its essential properties are lost in taking it apart.

Ackoff refers to elements *affecting* each other, which clarifies the implications of von Bertalanffy's *interaction*.

Von Bertalanffy defined an open system "as a system in exchange of matter with its environment" (1968b, p. 141). In accordance with Ackoff's approach, I would generalize this to "a system which is affected by any element or system not wholly contained within it". Note that according to Ackoff's principles, any effect on an element of a system, e.g., from the environment, affects the system.

The essence of GST's contribution has been to model *open* systems; traditional science implicitly assumed its objects of interest to be closed. This was not only exemplified by the experimental method, but by social science research which satisfied itself with the identification of *proximate* cause.

Empirically, there can be little doubt that human behaviour occurs within open systems as defined above. While there can be good *practical* reasons for excluding from consideration some level of suprasystems (or overlapping systems) from a particular research, awareness of the open nature of human systems should result in level of analysis being more appropriately selected than on the basis often taken of doing so precipitously and unconsciously, using preconceptions and first impressions, taking closure at a particular level as obviously manifest. The "real" is a *mystery* which can only be *approached* by the ingenious researcher who is not blinded by the obvious (Bachelard, 1928). Hence GST theorists consider the open system to be the general case; closedness rather than openness needs to be justified.

System Processes. The following discusses some of the most basic system processes.

- *feedback.* Information about a system's effect on the environment is fed back to the system; this can result in a change in system "direction". This is the basic concept for the cybernetic science of information and control, often demonstrated by the example of "self-guiding" missiles.

- *bifurcation.* This refers to the alteration of a "path" as a result of the introduction into a system of a new element. This concept is frequently found in chaos and process theory.

- *structure.* In the general case, system elements are not homogeneous; they are arranged into "parts" or subsystems, with different activities or "functions"; the quality, nature and patterns of interactions vary throughout the system. Structureless systems (e.g., gases) are special cases, whose modeling is a specialty of chaos theory.

- *2nd law of thermodynamics.* Entropy (disorganization of matter and diffusion of energy) has a tendency to increase, reaching equilibrium only at maximum entropy, resulting in a structureless system. However, entropy is inevitable only for *closed* systems, e.g. the universe (which is closed by definition), and does not apply to information. Hence the only system that can be "static" is a closed system in maximum entropy. A closed system not in maximum entropy is in a process of disintegration, while an open system is constantly being affected by its environment. The implication of the 2nd law combined with basic systems principles is that virtually all systems are *dynamic*: actively changing or adapting. Therefore, the fourth dimension is always explicit in systems analysis; systems can only be understood *over time*.

- *homeostasis* or *steady state.* Homeostatic systems maintain the same structure over time (*heterostatic* systems do not). Other than closed systems in maximum entropy, this can only be achieved by exchanging matter/energy or by interacting with the environment. The application of homeostatic analysis to organisms reveals the activity which the organism has to be engaged in, given the influence of the environment, in order to maintain the *appearance* of being unchanged. The application of homeostasis by von Bertalanffy and others revolutionized the field of biology, necessitating a shift in analysis from mechanics to *process*[13].

- *increasing complexification.* This is the idea that *open* systems have a tendency to develop increasingly complex structures; if not, they tend to disintegrate or disappear. In general, therefore, they are heterostatic. The tendency to complexification is a hypothesis based on observation. Many theories have been proposed to justify it, some of which are axiomatic but dependent on various assumptions about the nature of the system or its environment. Natural selection and other evolutionary theory is most common. Metaphysical propositions have also been advanced.

- *cycles, step theory, crisis, catastrophe, oscillation, etc.* Systems theory is developing (or adopting from disciplines such as physics) patterns and distinctive types of processes by which system change can be axiomatically deduced given various assumptions. Such processes have been clearly documented for various physical phenomena, and seem apparent in some processes involving information or social interaction. In general, the more complex the system and its environment, the less likely one is to observe or be justified in postulating precise patterns. Explicit use of systems theory can help to reduce the sloppy

---

13. The next step, of course, was to model organisms as heterostatic systems which include homeostatic functions, where the organism develops or changes gradually.

attribution of patterns to social/human processes, by suggesting to the researcher what assumptions or parameters would need to be satisfied.

- *self-organizing system.* Open systems reorganize themselves. This could be explained in a number of ways, depending on the nature of the system and on the theorist's predilections: in order to adapt to changes in the environment, defeat competitors, increase power, knowledge, happiness, to better achieve goals or to seek new goals, etc. Reorganization may involve increasing complexification, leading to emergence.

- *emergence.* This is a concept that by its very nature perhaps must be vague. In its most widespread use, it refers to the change of a system into a "higher level" system. This is ordinarily viewed as more of an event than as a continuous phenomenon: while complexification, or increasing control over or enclosure of the environment, or growing knowledge, may be gradual to the point of emergence, the new system is fundamentally different in kind than the old. At some critical point a system that has been changing suddenly has greatly increased capacity, awareness, power, changes the way in which it operates, or adopts new goals or paradigms. While in retrospect, or from the perspective of a higher systemic level, an emergence might be explainable, it is not predictable based on analysis of the prior system or of its components. This process has been used to characterize the development of consciousness, social organization and civilization. The concept of emergence describes an abstract phenomenon which seems to have empirical counterparts; explanation requires use of process or cybernetic control theory.

Levels and Hierarchies. In a seminal GST paper, Boulding (1968[1956]) suggested that the new field organize its work around a hierarchy of systems organized around degree of complexity, with appropriate application to empirical systems. As a working hypothesis he presented such a hierarchy (frameworks, clockworks, thermostat, open system, plant, animal, human, social organizations, transcendental). This approach emphasizes that while there are similarities among systems, complexity renders them different in kind. An implication would be that appropriation of computer or organismic models to describe human phenomena would have to be justified.

Some researchers, e.g., Miller, have concentrated on documenting cross-level similarities — for example, between biological organisms and social organizations (Tracy, 1993) — based on particular theories of system characteristics. Such work — some of which claims to be strictly grounded — runs the risk, as warned by von Bertalanffy (1968b), of substituting facile analogism for the painstaking dialectics of scientific method aimed at establishing homologisms among system types.

The above summarizes some of the most basic concepts in GST. I have adapted them from several authors[14], not finding a single recent source which succinctly describes the key concepts in a way that reflects mainstream GST thought. These concepts were listed roughly in order of decreasing consensus. That is, there is less agreement in the literature regarding definitions and theory of complexification and emergence than of system and homeostasis. Systems theory is still a young science; terminology, concepts and theory have changed over time and vary in the hands of the sub-fields and related disciplines[15].

New Directions in Systems Theory

Some indication of the branching of systems thinking in terms of systems specialties and applications has been given above. Much of systems work today in the disciplines consists of developing theory with empirical content appropriate to the concerns or perspectives of each discipline. Some of these will be described in part 2 below. In a general sense, in recent years increased emphasis has been given to the concepts of complexity, chaos, and process.

---

[14] Among them are Bahg (1990); Bateson (1972); von Bertalanffy (1968); Capra (1990); Fivaz (1993); Gao and Charlwood (1991); Katz and Kahn (1978); Nagel (1969[1956]); Schwarz (1991).

[15] According to Waelchli (1992):

> The idea of *one* comprehensive thesis or proposition (let alone eleven) agreed to by all General System thinkers borders on fantasy. An essential, perhaps defining, quality of the General System thinker is that he or she owns a particular and singular world view and, from it, an unremittingly personal way of ordering human knowledge and experience. The quest for new and apt ordering principles is a hallmark of the GST thinkers I have known (p. 4).

Complexity. The recognition that humans and their behaviour are complex, within a complex environment, has become widespread in the methodological literature. Some authors (e.g., Almeida-Filho, 1994; Bibeau, 1982) explicitly propose a systems framework in order to model multiple dimensional contexts and organize data.

Many others adopt an "interpretive" approach which aims to benefit from the richness of human research objects (referred to as "subjects") by implicating the researcher in a dialectical reiterative process of interviews and surveys, detailed analysis of the results, followed by further interviews, etc. This approach has the merits of discovery, of researcher openness to what the "subjects" say about themselves, their motives and their contexts. It delves more deeply into human situations than does the more traditional use of socio-demographic statistical analysis.

Insofar as such research lacks *a priori* theory or an analytical framework, however, it can suffer from some of the disadvantages of modernistic methodology mentioned above. Every researcher has subconscious preconceptions bound to influence what is seen and the way in which it is analyzed, but stipulation of prior theory renders such preconceptions more transparent to other researchers. Also, lack of prior theory reduces the generalizability of results[16], although the interpretive method increases the likelihood that at least the actions of the *sample* will be understood (in comparison to more superficial statistical research).

As will be elaborated in this paper, GST offers an analytical framework which is very flexible in the degree to which models can be specified *a priori*, while improving the chance that the research will find data to be meaningful (i.e., in a dynamic context) and generalizable (see Frischknecht and van Gigch, 1989). Ball (1977) has suggested an approach which appears to combine interpretation with "systemic perspectivism", "leading to an investigative sociology" (p. 1). In referring to the sociology of knowledge he sees an epistemology without a methodology, but "general systems theory emphasizes *conceptual rigor*,

---

[16] Since without theory, we have no reason to believe that what is true for one population, as limited by factors such as space and time, will be true for another. See Ayer (1956).

increasingly appropriate *precision*, and the construction of verifiable *models* whose 'meaning' is determined by their *usefulness* to the solution of our problems" (p. 6).

Chaos. The field of "chaos" is, at least in the ordinary meaning of the word, in chaos. In recent years concern with chaos has become very fashionable, perhaps because of visual representations of chaotic phenomena in the media[17]. There does not seem to be consensus about exactly what a chaotic phenomenon is or what chaos theory concerns. Chaos has been referred to as the "science of complexity" (see Gleick, 1987, p. 5), which is somewhat inapt: most non-linear dynamic systems are complex in their behaviour, but, when compared with other such systems, systems commonly described as chaotic are usually more probabilistically deterministic (given knowledge of intervening variables), with relatively homogeneous elements and relatively little structure (Kremyanskiy, 1969). The behaviour of such a chaotic system might be describable in mathematical expressions called "fractals".

Some writers have experimented with applications of chaos phenomena to social behaviour (Dobuzinskis, 1992; Richards, 1990) or have suggested the potential of adapting chaos methodology to statistical analysis of epidemiological phenomena (Almeida-Filho, 1994). Certainly the large amount of new research in this field promises enrichment of systems theory and of social science methodology. We need only to warn of the potential for indiscriminate chaos applications to again reduce social behaviour to probabilistic determinism and mathematics; in other words, to substitute technique for content. Systems philosophers emphasize human value diversity and increasing consciousness and intelligence. Emergence — self-organizing dynamic behaviour — in this context is not in any sense deterministic.

---

[17] Images of chaos are conveyed by numerous documentaries, television shows, magazine articles, and popular books, which use colourful photographs or film, and demonstrate the new capacities for computer generated graphics, modeling behaviour of macroeconomic, demographic and natural system behaviour. See, for example, the photos in Gleick (1987, pp. 114-115).

Process. It was perhaps inevitable that the increased concern with *interaction* of elements, within structures created by that interaction, would lead, in its most extreme form, to the dispensing of elements and structures altogether. While process and interactionist theories don't usually go to this extreme, they differ from traditional systems theory by reducing emphasis on structure, functions and roles, and elements (e.g., Chubb, 1990). Process theory has been receiving much attention throughout the sciences over the last decade[18]. As with chaos theory, process theories or approaches may be more valuable in elaborating, or animating, aspects of abstract system theory than in their possibly indiscriminate application within disciplinary theory[19].

Much of the future of GST will result from specification of *types* of systems or of particular system *behaviours*. Appropriate application of this research in the human sciences will need to bear in mind the limiting assumptions of such system models, and *neither substitute models nor motions for meaning*. I would prefer to see the systems school as building a *framework* for synthetic analysis, as opposed to producing a series of confectionery concoctions to be snapped up by consumers drooling for the latest fashionable flavour, resulting in inappropriate application of new developments such as process and chaos theory[20].

---

[18] However, process theories have long roots, predating and contributing to early GST development. See Buckley's literature review (1967, pp. 17-23); also Deutsch (1965[1953]); Leavitt (1965[1951]); and Lewin (1965[1947]).

[19] See Sabelli (1991), for a process theory based on postulates of asymmetry, opposition and bifurcation.

[20] In various places in this paper I refer to "inappropriate application": it should be understood that such references are not comments about the basic theory being applied, but of the use to which it is put. The long-term development of promising new theory is hampered when initial enthusiasm results in over-extension, tarnishing the entire theoretical approach.

Part II: The Influence Of GST in the Social Sciences

Infusions from the Pure and Natural Sciences

While social scientists (Boulding, Parsons, Buckley, Simon, Easton, Ackoff) made large contributions to the early growth and application of GST, many roots can be found in the theoretical work of the natural sciences, particularly biology, physics and engineering. In those fields advancement increasingly required extensive theory rather than pure observation (e.g. relativity theory, quantum mechanics, "discovery" of the D.N.A. structure). In a sense, science had to step back thousands of years and replenish itself with Socratic/Platonic concerns about dialectic, form and abstraction.

In biology, the realization that organs and organisms could only be understood within a larger context, i.e. the "organismic" conception advanced by von Bertalanffy (1968b), has led to new disciplines and approaches, some of which explicitly incorporate system theories[21]. In 1968 Buckley referred to the "many empirical studies in areas of biology demonstrating beyond doubt the fruitfulness of modern systems analysis" (p. xi). Köhler's (1947) application of systems concepts in developing his gestalt theories has many counterparts today in research in cognitive psychology as well as in neurology, some of it addressing issues of emergence and consciousness (e.g. Vandervert, 1991a and 1991b). Systems theory plays a major role in the new field of ecology, integrating animate with inanimate matter over time and space[22].

Physicists and engineers (Prigogine, Haken, Tsuchida, Eigen) have made important recent contributions to systems theory, dealing with thermodynamics, entropy, synergism, hypercycles, resource physics, and dissipative structures (Bahg, 1990). The exploding interest in chaos theory has resulted in cross-fertilization among virtually all the natural and mathematical sciences; Gleick stated that:

---

[21] See, e.g., Cascante, Sorribas, Franco, and Canela (1991); Miller and Miller (1992); Savageau (1991); Sharov (1991); Voit (1992).

[22] See: Auger (1991); Botkin (1976); Lovelock (1979); Wolanski (1989); Wu (1976).

> Chaos breaks across the lines that separate scientific disciplines. Because it is a science of the global nature of systems, it has brought together thinkers from fields that had been widely separated ... specialization has reversed because of chaos (1987, p. 5).

Systems theory has made the boundaries between the natural and human sciences increasingly fuzzy. Several universities today have "systems science" departments staffed by sociologists, physicists, psychologists, engineers, computer scientists, management researchers, biologists, policy scientists, and mathematicians, along with researchers who can only be classified as metadisciplinary, as system scientists, or as adisciplinary.

GST has influenced directly, and enormously influenced indirectly, most of the major social science fields. Direct influence can be seen in the fields of psychology, sociology and political science, in which writers have made seminal contributions to mainstream GST. These writers launched new research directions in their fields that represented a major divergence from the extant states of their disciplines. These contributions gave birth to approaches with new names, often merging systems parentage with streams of postmodern thinking and methodology.

Below I will give a rough summary of GST development and application among some social science disciplines; it should be borne in mind that given the inherent inter/metadisciplinary nature of GST, division of its social science applications among disciplines can be somewhat inappropriate[23].

Philosophy and Methodology

Michael Polanyi's (1967) work addresses the development of thought as a philosophical question within a context of culture and values. Rejecting causal determinism, he saw thought as "emerging" through the efforts of "explorers". Interestingly, he took a pragmatic view of what does and should constitute a successful theory or paradigm: acceptance by the community of scientists. Rather than being relativistic, the import of Polanyi's views is that whether or not science progresses depends on a *social* process; science does not propel itself, it is

---

[23] Applications of systems approaches in the policy sciences, health disciplines, and mental health field are reviewed in McCubbin (1994b).

propelled by the efforts of individuals in community. He incorporates basic GST ideas of gestalt, structure, hierarchy, and emergence in his description of the development of thought:

> Processes that are expected to achieve something have a value that is inexplicable in terms of processes having no such value. The logical impossibility of such explanation may be affiliated to the dictum that nothing that *ought* to be, can be determined by knowing what *is*.... Thus the logical structure of the hierarchy implies that a higher level can come into existence only through a process not manifest in the lower level, a process which thus qualifies as an emergence (pp. 44-45).

Kenneth Boulding is a founder of GST who not only helped to develop the basic concepts of the field (e.g., 1968[1956]) but has been a prolific writer in developing an application of GST now known as "world systems theory". Although an economist, he does not appear to have influenced mainstream economics very much with respect to systems theory, but his thought has had major impacts in general systems philosophy, political economy, ecology, and international development theory. While most GST applications take as their focus of study the individual, the group, the organization, the institution or the state, Boulding has made unique and important contributions in modeling the *world* as a multi-dimensional system (as an entry point into a large amount of world systems theory see Boulding, 1985).

T. Downing Bowler is an eminent GST thinker whose book *General Systems Thinking: Its Scope and Applicability* (1981) is a crystal-clear synthesis of systems thinking from a wide variety of natural and social science fields and mathematics, expressed with the precision of a professional philosopher. This book culminates in a very useful list of GST assumptions and of inferences drawn from the present state of knowledge, which provide the basis for GST "as a philosophy of world and life" (pp. 219-222).

> These assumptions and statements constitute a model of the nature of the universe and the conditions for any kind of existence. They are derived by beginning at the highest level of generalization, and provide a framework for research into the areas of specialized knowledge for verification and improvement of general applicability. They have heuristic value for problem solving, because they indicate relations that should be researched .... This model appears to be applicable to all systems, whether formal ... existential ... or affective. ... In this sense, the problem of ontology ... is replaced by statements concerning relations .... This model provides a context within which fragmented knowledge and education can be reunited, in the humanities, arts, sciences, and formal disciplines. Finally, it provides a framework within which

> humanity can locate itself and the world and within which the ancient human concerns can be redefined in more manageable terms.

John Sutherland's *A General Systems Philosophy for Social and Behavioral Sciences* (1973) provided an important milestone in the development of GST, in that it systematically examined the epistemological foundations and methodological implications of GST. This work injected some needed coherence into what had become a messy field (messy in terms of the wide variety of work associated, correctly or not, with GST), and remains indispensable to the researcher wishing to develop empirical applications of GST who is conscientious about the epistemological/methodological underpinnings of the GST approach. Sutherland attacks scientism while emphasizing the importance of empirical work in a process of deduction and induction, and is critical of perspectives overly constrained by disciplinary perspectives and theories:

> Thus we find the general systems theorist arguing for an end to academic parochialism and for adoption of interdisciplinary attack; arguing against simple statistical-mathematical models ... and asking for more elegant and relevant formulations .... because his commitment is not to any discipline or school, but to a philosophy ... freed from the exegetical methodology which 'schoolism' implies; and he is free from the necessity to pay homage to any academic abstractions such as Freudian psychology, functionalist anthropology, or Parsonian sociology. ... For when one is captured by a disciplinary dogma, one ceases to be scientist and becomes evangelist, ceases to be investigator and becomes concept-mongerer (pp. 189-190).

There is an enormous amount of work in developing methodologies, including modeling and empirical verification, most of which is highly technical and intended for particular types of applications. Some of this literature is inaccessible to the researcher lacking background in systems concepts and language, mathematics, or computer science, yet it contains quickly multiplying nuggets that could be profitably mined by the social science researcher seeking to appropriately model and test non-linear dynamic complexity. Peter Checkland's *Systems Thinking, Systems Practice* (1981) provides an easily read heavily used survey of systems methodologies, and develops his own problem-solving orientation. Jean-Louis Le Moigne is among the foremost systems methodologists; his *La théorie du système général: Théorie de la modélisation* (1984) is an important text both for general systems researchers and for operational systems work (see also Charest, 1980).

Political Science

Since David Easton's seminal work (1965), along with contributions by Karl Deutsch, Kenneth Boulding, Talcott Parsons, and Gabriel Almond, systems theory has so dominated political science (Dougherty and Pfaltzgraff Jr., 1971) that it has become a paradigm for the field; it has become ingrained to the point that it is rare to find titles or abstracts describing a work as "applying systems theory". Nevertheless, GST may have had decreasing influence in mainstream political science in recent years; older GST concepts have been appropriated, used and abused, without continuous enrichment from GST as that field further refines its concepts and methodologies.

As a result, while GST was largely responsible for the injection of some "hard" tools into political science (multivariate models, game and decision theory), displacing the field's prior emphasis on Marxism, verbal description and case studies (focused on the organs of the state rather than on political processes in general), in recent years an increasing component of political science research consists of theory-less quantitative research which is scientistic rather than scientific. Systems concepts (e.g., cycles) seem to be increasingly applied *ad hoc* and *ex post facto* to statistical analysis, rather than *a priori*. Aside from such work however, the continued influence of postmodern philosophers like Popper (1979), Foucault (1964, 1966) and MacIntyre (1973, 1977, 1983) have provided a counterweight to empty exercises in correlation.

The new interdisciplinary fields of public policy and administration and public choice, founded in particular by Simon (1969[1956]), Lasswell, and Buchanan (1972) remain heavily influenced by GST and associated streams, particularly game and decision theory based on rational choice. Outside these new research areas, a largely "postmodern" backlash against the use of rational choice and utility theory has been widespread in recent years in the social sciences. The critique of rational choice has justification where the assumption has been used in simplistic linear causal models, but to a large degree it is premised (e.g., Stewart, 1993) on the earliest conceptions of rational choice.

Within a systems approach "irrational" decision-making might be understood as rational in its contexts of values, goals, uncertainty, strategy, constraints and opportunity; the hypothesis of rationality can also be justified on pragmatic and epistemological grounds (McCubbin, 1997; see Granovetter, 1985; Sloep, 1993). Without some fundamental postulates about human behaviour, we may be left with nothing but *explication* of things that "just happen", enclosing decision processes within a "black box" — hardly a base for scientific advancement.

The field of international politics has been superseded to a significant extent by new interdisciplinary approaches (e.g. international studies, international relations); as such, since its inception it has provided very fertile ground for GST and game theory. Boulding's world system theory has been used to develop a new international political economy, to some extent displacing or absorbing Marxist thought.

Psychology

Despite the continued strength in the field of psychology of scientistic methodology, including vestiges of behaviourism, there is a variety of thriving contextual (Reese, 1991), ecological, transactional and process approaches that incorporate much of GST thinking, although the various schools have tended to distance themselves somewhat from their systems kinship. Arthur Koestler's challenge to behaviourism, notably in *The Ghost in the Machine* (1967), provides a significant — but eclectic — contribution to GST thinking, particularly in psychology and biology. It provides an "emergence" perspective on evolution, and argues that there is nothing pre-ordained about the continued survival of the human race; alteration of our own beings[24] may be necessary if humankind is to avoid self-destruction.

The work of the "Palo Alto School", especially by Gregory Bateson (1972) and Paul Watzlawick (1976), continues to reverberate, not only in psychology but

---

[24] Koestler advocated the development and use of drugs to alter the human brain and behaviour, advice which appears to have been increasingly heeded by society and, in particular, by the mental health system. This is not the place to discuss my own abhorrence of the chemical engineering of the human personality; Ivan Illich (1977), David Cohen and I (Cohen, 1996; Cohen and McCubbin, 1990; McCubbin, 1994a; McCubbin and Cohen, 1996) have discussed the political economic pressures involved in pushing drugs for mental health, resulting in irrational outcomes from a societal perspective.

throughout society. These writers merged GST with game theory (particularly "prisoner's dilemma" and other paradoxes), information theory, futurism, eastern philosophy, and postmodern epistemology, resulting in syntheses challenging to the spirit as well as to the mind. The Palo Alto School provided an ecological approach to the understanding of the human psyche and behaviour, inspiring numerous transactional/interactional theories and therapies aimed at raising the potential of humans and of humanity to change their own natures given the *will* to do so, the *belief* that they can, and the dynamic development of self-knowledge (including learning how to learn; see especially Bateson, 1972, pp. 271-339)[25].

Some interesting recent examples of the systematic application of systems theory to psychology are provided by Heylighen (1992; self-actualization) and Lester (1993; personality disunity). Joanne Wieland-Burston (1992) has written a book applying chaos theory to psychotherapy.

Sociology

Ironically, the massive early influence of systems theory in this field provided its own undoing. Talcott Parsons' structural/functionalist model (1951, 1974 [1957]) was a systems theory, but it lacked the dynamic interactionism of more mature GST. Parsons' approach dominated sociology for more than twenty years, until the mid-1970s. Despite the efforts of some writers to update sociological systems theory (Bailey, 1993 and 1994; Buckley, 1967; Crozier and Friedberg, 1977; Katz and Kahn, 1978), systems theory has been struggling in mainstream sociology. Criticism of systems approaches sometimes verges on the perverse, bemoaning the lack of dynamism, system openness, multi-dimensionality, or "agency" — all of which have been emphasized and developed in *mainstream* GST since the mid-1960s.

Another (related) major axis of criticism is the supposed organicist approach of GST (e.g., Katz and Kahn, 1978, pp. 8-9). This remains a valid criticism of the

---

[25] It was Bateson's *Steps to an Ecology of Mind* that turned me into an enthusiastic advocate of GST thinking in both academic applications and personal philosophy, after literature including the work of George Bernard Shaw, and Steinbeck's *Grapes of Wrath* (1939), had created fertile ground for systems thinking, teleology, and collective consciousness.

Miller "living systems theory" school (Swanson, 1992; Tracey, 1993; see Miller and Miller, 1990), and of those still adhering to Durkheimian conceptions, but insofar as it takes *applications* of GST within biology as representative of *general* systems theory, one could say that the early spectacular successes of GST theorists like von Bertalanffy hurt the image of GST in general[26].

Despite the distorted image of GST in mainstream sociology, the basic ideas of modern GST have been advanced most notably in what could loosely be defined as the organization theory field and particularly in its recent offshoot of institutional theory. The work of Crozier and Friedberg (1977) emphasizes human agency, whereby individuals make strategic decisions in an environment of constraints and uncertainty; *L'acteur et le système* made a major contribution to GST by de-emphasizing what had become overly reified ideas about structure, and by emphasizing the importance of decision-making under uncertainty, reflecting the rapid developments in decision theory over the previous five or ten years.

Hasenfeld (1983, 1986) developed a political economy of human service organizations that uses a framework very close to GST. Hasenfeld's approach has been very influential; the merging of a GST approach with a political economy *perspective* (see also Benson, 1975) shows how a GST researcher can create *grounded* theory for use in particular research areas. A political economist researcher using a GST approach is less likely to see variables of power and money as *determinant* or *exhaustive*, but rather as providing particularly important *dimensions* constraining and influencing human decision-making.

John Meyer and W. Richard Scott have helped to create the field of *institutional theory* as a subject within sociology (along with others such as Rowan, DiMaggio, Powell; see Scott, 1993). Meyer and Scott have systematically applied exceptionally broad cultural, temporal and spatial dimensions to organizations and to social institutions and practices in the more abstract sense. The import of their research is that social organizations today need to be understood *as constructions of the social environment* (Meyer and Scott, 1992). Meyer and Scott's

---

[26] The same sort of irony relates to a common public association of GST with cybernetics, such that GST is seen as technocratic; the successful application of cybernetics in engineering and computer science leads many to think of systems theory as concerned with machines.

ideas could have some provocative implications in an analysis of the mental health system:

> The schooling, health care, and welfare organizations we look at look less like rational organizational structures than holding companies incorporating various institutionally defined packages. They are less bounded from their environments than immersed in them, deriving from them legitimacy and meaning. .... [T]he institutionalization of rational organization tends to lower the formal rationality of specific organizations. *The question is whether formal organizations dominate society or whether society has, through the long process, come to dominate formal organization.* I will here consider historical processes by which rational organization, institutionalized, destroys itself and becomes another form (emphasis added; Meyer, 1992, p. 262).

Scott has written a rare explicit application of GST to the mental health sector as a whole (1986; discussed in McCubbin and Cohen, 1997, which provides the only other such application).

Robert Atchley's article "Retirement as a Social Institution" (1982) provides an excellent example of how social institutions can be understood as dynamic open systems. He brings in political, economic, ideological and structural factors in describing how "retirement" in the U.S. evolved from 1700 to today. The long historical perspective reveals the importance of cultural and ideological factors lacking in much social science "snapshot" or short time frame research; such research can restore to our view the *collectively negotiated* human agency which in short term analyses seems to be lost under the enormous pressures of the environment:

> This brief history of retirement was intended to provide a sense of the social processes that have influenced retirement. Social ideology has clearly played an important part at every turn in the evolution of retirement. Indeed, the history of retirement cannot be understood simply in terms of economic or demographic processes, although they have played important roles. Intergroup dialects have also played an important part .... Another important insight from the history of retirement is the growing separation of the concept of retirement from the concept of old age (p. 274).

After developing his "systems theory of social institutions" and evaluating alternative theories of retirement, Atchley concludes:

> The early theories faltered because they tried to tie retirement to abstract social structures or social processes. But retirement has evolved through negotiations among major economic and political interest groups in America. These negotiations took place in the context of competing values and ideologies. .... The future of retirement as a social institution thus depends on who the groups are in the negotiation process, how powerful they are, what ideologies they pursue, and what practical economic, demographic, or social constraints they face. If this paper has helped to identify the

> actors, the goals they might want to pursue through retirement, and the major constraints they must negotiate within, then it has contributed to the development of an adequate theory of retirement. What remains is the task of carefully observing the negotiation process. This is not easy because retirement is a complex institution that like so many American institutions is both centralized and decentralized. (pp. 285-286)

Atchley's paper would provide an exemplary approach for the GST researcher developing a *problématique* for understanding the mental health sector as a dynamic complex open system, propelled by individual and collective value-based decisions made within changing historical power-structured contexts[27].

Part III: Hopes and Limits for GST in the Social Sciences

Limitations and Critiques of GST and Applications

Most criticism of GST is in fact criticism of *applications* to empirical situations. Researchers employing systems approaches should recognize that their work does not start with a given system model, but with creating a model based on both assumptions and observations — which should be made explicit. It is very easy to find data supporting analogies in cross-level or inter-system comparison; to become more scientific comparisons should be based on *a priori* selection of comparison criteria and of the means for operationalizing them.

To some degree system sub-fields have lost the degree to communicate with each other; avoidance of this would be enhanced by the sub-fields recognizing GST as the metascience of the systems field, using GST forums as the place to settle fundamental questions of theory and definition, and by rendering explicit what types of systems the sub-fields deal with. GST and its applications run the risk of any theoretical approach of over-protectiveness, resulting in stagnation and over-extension (e.g., Living Systems Theory). After GST is introduced in a discipline

---

[27] E.g., McCubbin, 1994a. Factors which constrain choices can be abstractly summarized as creating power structures; these can be anything from legal authority, economic means and disparities, violence, influence, as well as natural factors like disease, weather, geology, etc. Of course, power-structuring by its very nature creates and shapes opportunity.

and applied in creating a new disciplinary research program, it should receive constant rejuvenation from developments in mainstream GST and other metadisciplines, including epistemology, in order to avoid stagnation — and subsequently tarnishing the reputation of the systems approach as a whole.

Some key concepts, e.g. chaos, dynamism, complexification, and emergence are subject to a degree of controversy and ambiguity. This is natural and probably healthy for the growth of a young science; nevertheless, GST researchers and those who use those concepts should make an effort to ensure that they are well-explained or are consistent with mainstream GST use. The future of systems applications is hurt when researchers convey misleading ideas to their readers and each other.

In many respects the success of GST can hurt it. It is quite possible that conscientious application of GST to a real-world situation will reveal that system processes are so multiply interactive that we cannot adequately model it or make predictions. This might scare off some researchers who would rather have a wrong answer than none at all, or who unreasonably cut corners. My view shares that of Ackoff (1974), that it is better to have a poor solution that improves than a good one which doesn't; there is a need to begin creating rudimentary models of complex social phenomena, being honest about their limitations, in the hope that they can gradually be elaborated and tested. Consistent with the idea of emergence, we should take the long view, not knowing what new insights or methodologies might in the future transform our models.

Researchers eager to jump on and overextend the latest systems-associated trends and concepts — a danger evident with chaos and process theories — should recognize that their work could quickly become outdated. While explicit GST work may not always be fashionable, it has a chance to be understood by other GST researchers, provide a basis for further development in the field, and achieve the longevity likely to be denied to the fashions.

Economics has demonstrated how the desire for methodological clarity, or to appear scientific, through the use of formal logic, mathematics, and mathematical statistics, can push all other considerations aside. While GST applications can benefit from mathematical expression, particularly within particular stages of the research process as a means to improve the clarity of concepts and deductively generate unexpected hypotheses, the field needs to ensure that its work is communicable and that it does not base its research questions on technique or what is measurable.

Some systems applications (e.g., of Parsons) have been criticized for over-emphasis on structure, for being too closed, for being "mechanicist" (McLean, 1986) or "organicist" (Low, 1982), or for modeling equilibrium or homeostasis rather than heterostasis. I suspect some of that criticism is superficial, not based on a full understanding of the models, but some of it is certainly justifiable.

It is irritating to see, however, how often all of systems theory is tarred with the same brush, or how often critiques of a field of systems application ignore advances over several decades, when the specific contributions critiqued were advances when they were put forward, and when the research field has already satisfied the critiques (with respect to structural/functionalism, for example, see the work of Hasenfeld and of the institutional theorists — work that *builds on* the tradition established by Parsons and his contemporaries). A related point should be noted: any attempt at modeling a complex reality "goes on a limb"; it is liable to make assumptions which are clearly not fully reflective of the reality being *modeled*. The modeler should not apologize for this, but make clear what corners were cut and why.

There have been suggestions that GST, or certain applications, "reify" systems, boundaries, and other structures (Pam, 1993). I have yet to read or hear a convincing example of such criticism. It is hard to imagine a coherent theory or perception which doesn't rely upon structuring its objects through the use of signs and signifiers. In my view, whether "reality" itself contains structure is irrelevant to the business of science, since we can never "know" it as that of reality rather than as that of our appreciation of reality. The scientific endeavour consists of taking relatively amorphous messes and making "sense" of them — or, in the sense of Bachelard (1969), adding meaning to the "obvious" by reinterpreting it.

GST shares with most postmodern schools — e.g. constructivism (see von Glasersfeld, 1988) — the explicit recognition of the abstract nature of what is "observed". GST is primarily an epistemology and makes no claims about reality. Appropriate applications of GST are not statements of fact but rather hypotheses to be tested, exemplified by the paper by Revans on hospitals (1990[1962]) and Russell on hospices (1989). The favourable results of such tests can, at most, "support" our hypotheses — or, rather, we choose not to reject them — rendering the theory from which they are derived reasonable within an extensive framework of epistemological and methodological assumptions.

It is possible that some GST researchers do treat their systems/boundaries as self-evidently concrete; however this error seems most commonly made by *non-GST* researchers — e.g. in the administrative and policy sciences — who have shared in the wide appropriation of GST terms and concepts, like "system" and "subsystem", "feedback", and "dynamic", without the benefit of GST epistemology[28].

Studying the object of interest as an "open system" is inadequate, because this method cannot adequately account for the system's dynamic or developmental behaviour (McCubbin and Cohen, 1997), a point also made as part of a careful argument by Pondy and Mitroff (1979). Rather, the *object of study* should be a level above, or abstracted from, the *object of interest*[29]. This might imply the irony of increasing interdisciplinarity in order to better "specialize" with respect to the object of interest.

The synthetic approach in general suggests that excessive analysis of the object of interest is sub-optimal (given time constraints). Hence, for example, the GST researcher choosing the mental health system as object of interest may have to sacrifice the reading of a portion of mental health system literature in favour of

---

[28] I may not have provided full justice to this critique, however, due to my postpositivist inability to conceive of describing "reality" without structuring it (see Fischer, 1993, for a discussion of postpositivism in the policy sciences). It is an apparently unresolvable problem in the justification of a science that any justifying argument is itself made pursuant to a theoretical paradigm: the implied process of recursive justifications is infinite.

Radical relativism is one response to this problem; my response is based on the view that any response, including relativism — other than a nausea (in the sense of Sartre) so profound that it becomes personality disintegration, coma, or death — implicitly or explicitly relies on assumptions of belief, values, and faith, so one might as well choose "good" values and pragmatic beliefs (see James, 1970; McCubbin, 1997).

[29] Although there is a trade-off, in that this carries with it the risk of missing important aspects of the system of interest. Braverman (1993) suggests that family systems therapy has lost sight of the individual within the system, particularly with respect to the development of intimacy.

suprasystemic literature like philosophy, economics and social psychology. In general the GST researcher needs to be aware of the imperative ensuing from GST epistemology to consciously separate object of study from object of interest.

It is not necessarily the case that GST provides a great advantage in comparison with other epistemological approaches (e.g., provided by critical thinkers like Hegel and Foucault) which focus on the highest systemic levels (including historical) as part of a synthetic study of lower-level systems. However, systemic meta-epistemology suggests that "suprasystems" can be defined not only by space, time, and other empirical criteria, but also by level of abstraction, to which there are no limits (Overton, 1991a, 1991b). GST epistemology could be enriched by the research in the "archaeology of thought" by writers notably including Foucault (1962, 1975)[30] and Deleuze (1988).

Systems Thinking as a Bridge between Ethics and Science

Systems theory has had a great deal of influence throughout the social sciences, often mutating into disciplinary variants. In some fields, notably international politics, family therapy, and institutional theory, GST and its offshoots continue to develop with infusions from the main body of GST literature and from disciplinary theory and observation. Some fields such as the political and policy sciences have benefited from GST concepts applied decades ago, but show little further development and have not maintained bridges to the metadiscipline of GST. In sociology the enormous early influence of functional/structuralism came under such a widespread critique that overt systems approaches were reduced to voices crying in the wilderness, despite the excellent contributions from writers such as Buckley, Hasenfeld, Katz and Kahn, and Scott, whose social system theories went a long way to responding to the criticisms regarding process, agency, organicism, and dynamism.

The process of GST diffusion seems to have continued, but at a slower pace than that seen in the 1960s and early 1970s. It is to be expected that in a culture of

---

[30] A few systems applications, in what may well be a trend, have substantively drawn on the thought of Foucault (Flaskas and Humphreys, 1993), whose work has only recently begun entering the mainstream of English-speaking North American academia.

disposability novel approaches will initially be greatly exploited, often superficially, until the inevitable stream of critiques arrive, at which time the grazers move onto another fashion. The development of GST may also have been hurt by the right-wing entrenchment seen in western countries since the 1970s: GST not only challenges modernism but encourages explicit attention to suprasystems, leading to recognition that power and economic privilege, rather than rationality, justice, and democratic ideals, often underpin the status quo.

Part of the survival ability in a conservative era of some of the more relativist "postmodern" or sociology of knowledge approaches may be due to them being relatively *safe*, insofar as attention is directed exclusively to social constructions of problems, narrative, and dialectic, *as the objects of study and interest*. An explicitly value-based or theory driven stance can drive the researcher with a sense of outrage about injustice, or lead the researcher to systemic preconditions and constraints affecting what people experience and what choices they may make, rather than what is being talked about in a media and literature which is itself subject to such systemic contexts[31].

Ironically, the acceptance of GST has been hurt by widespread misperceptions that in effect equate it with extensions of modernism like linear programming, cybernetic studies of closed value-free systems, and management "systems analysis" aimed at more effectively implementing Taylorian ideals. It needs to be stressed that *General* Systems Theory is not synonymous with applications incorporating assumptions of process linearity, equilibrium/ homeostasis, closedness, deterministic causality, etc. As is true for many promising concepts and paradigms — e.g. socialism/communism, Christianity, liberalism, "community" — GST could be drawn upon to oppress as well as to liberate. It should be noted, however, that all of these modes of thinking contain some fundamental liberating ideas which we might expect a collective intelligence to periodically return to, ponder, and further develop, despite periodic distortion and manipulation by the few.

The application of spiritual or ethical paradigms can raise the systemic level of research (e.g., Hegel, Feuerbach, Fromm). Furthermore, while I cannot see GST

---

31. See Rosenau (1994) for an interesting discussion of the implications of postmodernism for health politics and for conduct of health politics research.

as *inherently* ethical, use of its concepts greatly facilitates the understanding and expression of ideas about humanity and the universe that underlie certain types of ethical stance. The "interconnectedness" of our universe suggests enhancing our ideas about social and ecological responsibility: the Kantian "enlightened self-interest". Our understanding of the capacity of complex systems to organize themselves *and* their environment suggests teleology rather than causal determinism, reinforcing not only our free will but our *belief* in it: liberating us but also driving home our responsibility to ourselves and others. Methodologies for simplifying complexity and developing strategy in a dynamic uncertain world offer at least the hope that through rational action we might be able to progress toward the "good" that increased awareness and responsibility suggest.

The perception of how individuals are enmeshed in social systems which are enmeshed in historical and ecological systems contributes to a blurring of atomistic self and national identities. Indeed, some of the spiritual implications of *universal* systemness become obvious upon reflection. Concepts of "intelligence" based on complex information processes can be applied not only to the "brain" but also to collectivities — and perhaps also to ecologies, cosmologies, or even the universe. It is but a short step to enriching with systems and information theory concepts of "alienation" handed down by Hegel, Feuerbach (1957), and Marx, to see a collective intelligence as alienated within itself insofar as its information processes are fragmented and distorting, with subsystems consequently experiencing isolation.

The point of GST is not to create a "picture" of reality, parceling it out into sets of objects, but to animate it. GST, along with some of the less cynical and relativistic postmodern approaches, hopes to replace black and white lithographs of our universe with four dimensional *interactive* images complete with colour and motion over historical time[32]. The abstract attribution of boundaries and hierarchies is not an end in itself, but rather a means to facilitate the application of

---

[32] Perhaps we could view the application to these images of critical self-consciousness — trying to make sense of our processes of making sense — as the fifth dimension of scientific enquiry.

theory about the *actions* of "systems" that can include anything from individuals to societies to schools of dolphins to snowflakes[33].

That GST envisages the application of systems thinking to thinking itself helps to clarify and implement ideas expressed by non-GST critical thinkers about levels, creases, dimensions, or folds, of abstraction, complexity, or perception. GST can be thought of as a meta-epistemology in that it offers a way to organize thought not only about the real world, but about thought itself, its development, its meaning, and its value.

---

[33] Despite the enormous complexity of the snowbank, containing billions of *unique* snowflakes, by incorporating levels of abstraction we can create "simplifying" theory about why snowflakes are different, and about how quickly the snow under a footprint will melt compared to untouched snow in the footprint's environment. By simplifying the complexity of snow, we can learn how to shovel a pathway through it.

References

Ackoff, R.L. (1974). The systems revolution. *Long Range Planning, 7,* 2-20.

Almeida-Filhõ, N. de (1994). *Towards a new paradigm in epidemiology: Ten drafted indications (...).* Unpublished.

Atchley, R.C. (1982). Retirement as a social institution. *Annual Review of Sociology, 8,* 263-287.

Auger, P.M. (1991). Influence of activity sequences on predator-prey models. In S.C. Holmberg and K. Samuelson (Eds.), *Systems science in the 21st century; integrating the new sciences of complexity in service of humans and their environment. Vol. I: Advances in natural systems science. Advances in informatics, communications and systems methodology,* pp. 44-50. Östersund, Sweden: Östersund University.

Ayer, A.J. (1956). *The problem of knowledge.* Baltimore, Maryland: Penguin.

Bachelard, G. (1928). *Essai sur la connaissance approchée.* Paris: J. Vrin.

Bachelard, G. (1969). *La formation de l'esprit scientifique: Contribution à une psychanalyse de la connaissance objective.* Paris: J. Vrin.

Bahg, Chang-Gen (1990). Major systems theories throughout the world. *Behavioral Science, 35,* 79-107.

Bailey, K.D. (1993). Prescience or serendipity? Parallelism in living systems theory and modern sociological theory. *Behavioral Science, 38,* 241-254.

Bailey, K.D. (1994). Talcott Parsons, social entropy theory, and living systems theory. *Behavioral Science, 39,* 25-45.

Ball, R.A. (1977). Equitable evaluation through investigative sociology. *Sociological Focus, 10,* 1-14.

Bateson, G. (1972). *Steps to an ecology of mind.* New York: Ballantine.

Benson, J.K. (1975). The interorganizational network as a political economy. *Administrative Science Quarterly, 20,* 229-249.

Bertalanffy, L. von (1968a). General system theory — A critical review. In W. Buckley (Ed.), *Modern systems research for the behavioral scientist: A sourcebook,* pp. 11-30. Chicago: Aldine.

Bertalanffy, L. von (1968b). *General system theory: Foundations, development, applications.* New York: Braziller.

Bibeau, G. (1982). A systems approach to Ngbandi medicine. In P.S. Yoder (Ed.), *African health and healing systems: Proceedings of a symposium,* pp. 43-84. Los Angeles: Crossroads.

Bibeau, G. (1988). A step toward thick thinking: From webs of significance to connections across dimensions. *Medical Anthropology Quarterly, 4,* 402-416.

Botkin, D.B. (1976). Population dynamics of long-lived animals. In J.D. White (Ed.), *General systems theorizing: An assessment and prospects for the future*, pp. 211-213. Washington, D.C.: Society for General Systems Research.

Boulding, K.E. (1968[1956]). General systems theory — the skeleton of science. In W. Buckley (Ed.), *Modern systems research for the behavioral scientist*, pp. 3-10. Chicago: Aldine.

Boulding, K.E. (1985). *The world as a total system*. Beverly Hills: Sage.

Bowler, T.D. (1981). *General systems thinking: Its scope and applicability*. New York: North Holland.

Braverman, S. (1993). The limits of systems therapy: The problem of intimacy. *Contemporary Family Therapy, 15*, 285-297.

Buchanan, J.M. (1972). Toward analysis of closed behavioral systems. In J.M. Buchanan and R.D. Tollison (Eds.), *Theory of public choice: Political applications of economics*, pp. 11-23. Ann Arbor: University of Michigan Press.

Buckley, W. (1967). *Sociology and modern systems theory*. Eaglewood Cliffs, New Jersey: Prentice-Hall.

Buckley, W. (1968). Preface. In W. Buckley (Ed.), *Modern systems research for the behavioral scientist: A sourcebook*, pp. xi-xii. Chicago: Aldine.

Campbell, J. (1982). *Grammatical man: Information, entropy, language, and life*. New York: Simon & Schuster.

Capra, F. (1990, summer). Life as a mental process. *Quest*, 7-11.

Cascante, M., Sorribas, A., Franco, R., and Canela, E.I. (1991). Biochemical systems theory: Increasing predictive power by using second-order derivatives measurements. *Journal of Theoretical Biology, 149*, 521-35.

Charest, J. (1980). *La conception des systèmes: Une théorie, une méthode*. Chicoutimi: Gaëtan Morin.

Checkland, P. (1981). *Systems thinking, systems practice*. Toronto: Wiley.

Chubb, H. (1990). Looking at systems as process. *Family Process, 29*, 169-175.

Cohen, D. (1996). Les "nouveaux" médicaments de l'esprit: Marche avant vers le passé? *Sociologie et Sociétés, 18*, 17-33.

Cohen, D., and McCubbin, M. (1990). The political economy of tardive dyskinesia: Asymmetries in power and responsibility. *Journal of Mind and Behavior, 11*, 465-488.

Collen, A., Minati, G., and Ciapessoni, E. (1994). Correspondence: Logical openness in systems. *Systems Research, 11*, 65-72.

Contractor, N.S., and Seibold, D.R. (1993). Theoretical frameworks for the study of structuring processes in group decision support systems: Adaptive structuration theory and self-organization systems theory. *Human Communication Research, 19*, 528-563.

Crozier, M., and Friedberg, E. (1977). *L'acteur et le système: Les contraintes de l'action collective.* Paris: Seuil.

Deleuze, G. (1988). *Le pli: Leibniz et le Baroque.* Paris: Minuit.

Deutsch, M. (1965[1953]). The effects of cooperation and competition upon group process. In H. Proshansky and B. Seidenberg (Eds.), *Basic studies in social psychology,* pp. 552-576. Toronto: Holt, Rinehart and Winston.

Dobuzinskis, L. (1992). Modernist and postmodernist metaphors of the policy process: Control and stability vs. chaos and reflexive understanding. *Policy Sciences, 25,* 355-380.

Dougherty, J.E., and Pfaltzgraff, Jr., R.L. (1971). *Contending theories of international relations.* Toronto: Lippincott.

Durkheim, E. (1981). *Les règles de la méthode sociologique.* Paris: Presses universitaires de France.

Easton, D. (1965). *A systems analysis of political life.* New York: Wiley.

Eckler, I.S. (1994, April). About today. *Psycho Magazine: Queen Street Patients Voice, 1,* 4.

Feuerbach, L. (1957). *Essence of Christianity.* New York: Harper.

Fischer, F. (1993). Reconstructing policy analysis: A postpositivist perspective. *Policy Sciences, 25,* 333-339.

Fivaz, R. (1993). Morphodynamics: Ergodic theory of complex systems. *Systems Research, 10,* 27-51.

Flaskas, C., and Humphreys, C. (1993). Theorizing about power: Intersecting the ideas of Foucault with the "problem" of power in family therapy. *Family Process, 32,* 35-47.

Foucault, M. (1962). *Maladie mentale et psychologie.* Paris: Presses universitaires de France.

Foucault, M. (1964). *Folie et déraison; histoire de la folie à l'âge classique.* Paris: Union Générale d'Éditions.

Foucault, M. (1966). *Les mots et les choses.* Paris: Gallimard.

Foucault, M. (1975). *Naissance de la clinique.* Paris: Presses universitaires de Frances.

Frischknecht, F., and van Gigch, J.P. (1989). The information-system-processing triad: Common core of human disciplines. *Behavioral Science, 34,* 239-252.

Gao, S.-J., and Charlwood, F.J. (1991). Emergence, structural stability and dynamical bifurcation. In S.C. Holmberg and K. Samuelson (Eds.), *Systems science in the 21st century; integrating the new sciences of complexity in service of humans and their environment. Vol. I: Advances in natural systems science. Advances in informatics, communications and systems methodology,* pp. 97-111. Östersund, Sweden: Östersund University.

Glasersfeld, E. von (1988). Introduction à un constructivisme radical. In P. Watzlawick (Ed.), *L'invention de la réalité. Contributions au constructivisme* (trans. A.-L. Hacker), pp. 19-43. Paris: Seuil.

Gleick, J. (1987). *Chaos: Making a new science.* New York: Cambridge University Press.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91,* 481-510.

Gray, W., and Rizzo, N.D. (1969). History and development of general systems theory. In W. Gray, F.J. Duhl, and N.D. Rizzo (Eds.), *General systems theory and psychiatry,* pp. 7-31. Boston: Little, Brown.

Hacking, I. (1990). *The taming of chance.* New York: Cambridge University Press.

Hasenfeld, Y. (1983). *Human service organizations.* Eaglewood Cliffs, New Jersey: Prentice-Hall.

Hasenfeld, Y. (1986). Community mental health centers as human service organizations. In W.R. Scott and B.L. Black (Eds.), *The organization of mental health services: Societal and community systems,* pp. 133-146. Beverly Hills: Sage.

Heylighen, F. (1992). A cognitive-systemic reconstruction of Maslow's theory of self-actualization. *Behavioral Science, 37,* 39-58.

Hofferbert, R.I., and Schäfer, G.F. (1982). The application of general systems methodology to the comparative study of public policy. *International Journal of General Systems, 8,* 93-107.

Illich, I. (1977). *Limits to medicine. Medical nemesis: The expropriation of health.* Harmondsworth, England: Penguin.

James, W. (1970). Belief and perception of reality. In H.S. Thayer (Ed.), *Pragmatism: The classic writings,* pp. 157-174. Toronto: Mentor.

Kant, I. (1958). *Critique of pure reason* (trans. and intro. by N.K. Smith; abridged). New York: Modern Library.

Katz, D., and Kahn, R.L. (1978). *The social psychology of organizations* (2nd ed.). Toronto: Wiley.

Kerlinger, F.N. (1986). *Foundations of behavioral research* (3rd ed.). Toronto: Holt, Rinehart and Winston.

Koestler, A. (1967). *The ghost in the machine.* London, U.K.: Pan.

Köhler, W. (1947). *Gestalt psychology.* Toronto: Mentor.

Kremyanskiy, V.I. (1969). Certain particularities of organisms as a "system" from the point of view of physics, cybernetics, and biology. In F.E. Emery (Ed.), *Systems thinking: Selected readings,* pp. 125-146. Baltimore, Maryland: Penguin.

Le Moigne, J.-L. (1984). *La théorie du système général: Théorie de la modélisation.* Paris: Presses universitaires de France.

Leavitt, H.J. (1965[1951]). Some effects of certain communication patterns on group performance. In H. Proshansky and B. Seidenberg (Eds.), *Basic studies in social psychology,* pp. 577-592. Toronto: Holt, Rinehart and Winston.

Lester, D. (1993). On the disunity of the self: A systems theory of personality. *Current Psychology, 12,* 312-325.

Lewin, K. (1965[1947]). Group decision and social change. In H. Proshansky and B. Seidenberg (Eds.), *Basic studies in social psychology,* pp. 423-437. Toronto: Holt, Rinehart and Winston.

Lovelock, J.E. (1979). *Gaia: A new look at life on earth.* Oxford: Oxford University Press.

Low, N. (1982). Beyond general systems theory: A constructivist perspective. *Urban Studies, 19,* 221-233.

MacIntyre, A. (1973). Ideology, social science, and revolution. *Comparative Politics, 5,* 321-342.

MacIntyre, A. (1977). Epistemological crises, dramatic narrative and the philosophy of science. *The Monist, 60,* 453-472.

MacIntyre, A. (1983). The indispensability of political theory. In D. Miller and L. Siedentop (Eds.), *The nature of political theory,* pp. 17-33. London: Oxford University Press.

McCubbin, M. (1994a). Deinstitutionalization: The illusion of disillusion. *Journal of Mind and Behavior, 15,* 35-53.

McCubbin, M. (1994b). *A systems theory perspective of mental health services and policy formation.* Unpublished monograph.

McCubbin, M. (1997). Free will, rationality, and explanation: Toward a teleological *human* science. Submitted to *Journal of Mind and Behavior.*

McCubbin, M., and Cohen, D. (1996). Extremely unbalanced: Interest divergence and power disparities between clients and psychiatry. *International Journal of Law and Psychiatry, 19,* 1-25.

McCubbin, M., and Cohen, D. (1997). Mental health policy from a systemic perspective: Orientations for strategic reform. Submitted to *Journal of Health Politics, Policy and Law.*

McLean, A. (1986). Family therapy workshops in the United States: Potential abuses in the production of therapy in an advanced capitalist society. *Social Science and Medicine, 23,* 179-189.

Mesarovic, M.D., and Takahara, Y. (1975). *General systems theory: Mathematical foundations.* New York: Academic Press.

Meyer, J.W. (1992). Conclusion: Institutionalization and the rationality of formal organizational structure. In J.W. Meyer and W.R. Scott (Eds.), *Organizational environments: Ritual and rationality* (Updated Edition), pp. 261-282. Newbury Park, California: Sage.

Meyer, J.W., and Scott, W.R. (Eds.) (1992). *Organizational environments: Ritual and rationality* (Updated Edition). Newbury Park, California: Sage.

Miller, J.G., and Miller, J.L. (1990). Introduction: The nature of living systems. *Behavioral Science, 35,* 157-163.

Miller, J. L. and Miller, J.G. (1992). Greater than the sum of its parts. I. Subsystems which process both matter-energy and information. *Behavioral Science, 37,* 1-38.

Morin, E. (1990). *Introduction à la pensée complexe.* Paris: ESF Éditeur.

Nagel, E. (1969[1956]). A formalization of functionalism. In F.E. Emery (Ed.), *Systems thinking: Selected readings,* pp. 297-329. Middlesex, England: Penguin.

Overton, W.F. (1991a). The structure of developmental theory. *Advances in Child Development and Behavior, 23,* 1-37.

Overton, W.F. (1991b). Metaphor, recursive systems, and paradox in science and developmental theory. *Advances in Child Development and Behavior, 23,* 59-71.

Pam, A. (1993). Family systems theory: A critical view. *New Ideas in Psychology, 11,* 77-94.

Parsons, T. (1951). *The social system.* Glencoe, Illinois: Free Press.

Parsons, T. (1974[1957]). The mental hospital as a type of organization. In Y. Hasenfeld and R.A. English (Eds.), *Human service organizations: A book of readings,* pp. 72-97. Ann Arbor: University of Michigan Press.

Polanyi, M. (1967). *The tacit dimension.* Garden City, New York: Anchor.

Pondy, L.R., and Mitroff, I.I. (1979). Beyond open system models of organization. *Research in Organizational Behavior, 1,* 3-39.

Popper, K.R. (1979). *Objective knowledge: An evolutionary approach.* Oxford: University Press.

Rapoport, A., and Chammah, A. (1965). *Prisoner's dilemma.* Ann Arbor, Michigan: University of Michigan Press.

Reese, H.W. (1991). Contextualism and developmental psychology. *Advances in Child Development and Behavior,* 23, 187-230.

Revans, R.W. (1990[1962]). The hospital as a human system. *Behavioral Science, 35,* 108-114.

Richards, D. (1990). Is strategic decision making chaotic? *Behavioral Science, 35,* 219-232.

Rosenau, P.V. (1994). Health politics meets post-modernism: Its meaning and implications for community health organizing. *Journal of Health Politics, Policy and Law, 19,* 303-333.

Russell, G. (1989). Hospice programs and the hospice movement: An investigation based on general systems theory. *Hospice Journal, 5,* 23-49.

Sabelli, H.C. (1991). Process theory, a general theory of natural and human systems. In S.C. Holmberg and K. Samuelson (Eds.), *Systems science in the 21st century: Integrating the new sciences of complexity in service of humans and their environment. Volume III: Advances in systems management and applications*, pp. 168-174. Östersund, Sweden: Östersund University.

Sahlins, M.D. (1965). On the sociology of primitive exchange. In M. Banton (Ed.), *Relevance of models for social anthropology*, pp. 139-236. Toronto: Tavistock.

Savageau, M.A. (1991). Biochemical systems theory: Operational differences among variant representations and their significance. *Journal of Theoretical Biology, 151*, 509-30.

Schwarz, E. (1991). From thermodynamics to consciousness. A model for evolution. In S.C. Holmberg and K. Samuelson (Eds.), *Systems science in the 21st century; integrating the new sciences of complexity in service of humans and their environment. Vol. I: Advances in natural systems science. Advances in informatics, communications and systems methodology*, pp. 235-241. Östersund, Sweden: Östersund University.

Scott, W.R. (1993). Recent developments in organizational sociology. *Acta Sociologica, 36*, 63-68.

Shakun, M.F. (1981). Policy making and meaning as design of purposeful systems. *International Journal of General Systems, 7*, 235-251.

Sharov, A.A. (1991). Self-reproducing systems: Structure, niche relations and evolution. *Biosystems, 25*, 237-249.

Simon, H.A. (1969[1956]). Rational choice and the structure of the environment. In F.E. Emery (Ed.), *Systems thinking: Selected readings*, pp. 214-229. Harmondsworth, Middlesex, England: Penguin Books.

Sloep, P.B. (1993). Methodology revitalized? *British Journal of the Philosophy of Science, 44*, 231-249.

Steinbeck, J. (1939). *The grapes of wrath*. New York: Modern Library.

Stewart, J. (1993). Rational choice theory, public policy and the liberal state. *Policy Sciences, 26*, 317-330.

Sutherland, J.W. (1973). *A general systems philosophy for the social and behavioral sciences*. New York: Braziller.

Swanson, D.L. (1992). Dysfunctional conglomerates: An explanation provided by linking ontological individualism to social relationships within an open systems context. *Behavioral Science, 37*, 139-151.

Tracy, L. (1993). Applications of living systems theory to the study of management and organizational behavior. *Behavioral Science, 38*, 218-230.

van Gigch, J.P. (1990). Systems science, the discipline of epistemological domains, contributes to the design of the intelligent global web. *Behavioral Science, 35*, 122-137.

Vandervert, L.R. (1991a). A measurable and testable brain-based emergent interactionism: An alternative to Sperry's mentalist emergent interactionism. *Journal of Mind and Behavior, 12,* 201-219.

Vandervert, L.R. (1991b). On the modeling of emergent interaction: Which will it be, the laws of thermodynamics, or Sperry's "wheel" in the subcircuitry? *Journal of Mind and Behavior, 12,* 535-540.

Varela, F.J. (1988). Le cercle créatif: Esquisses pour une histoire naturelle de la circularité. In P. Watzlawick (Ed.), *L'invention de la réalité: Contributions au constructivisme* (trans. A.-L. Hacker), pp. 329-345. Paris: Seuil.

Voit, E.O. (1992). Symmetries of S-systems. *Mathematical Biosciences, 109,* 19-37.

Waelchli, F. (1992). Eleven theses of general systems theory (GST). *Systems Research, 9,* 3-8.

Watzlawick, P. (1976). *How real is real? Confusion — Disinformation — Communication.* New York: Random House.

Wieland-Burston, J. (1992). *Chaos and order in the world of the psyche.* New York: Routledge.

Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine.* Cambridge, Massachusetts: M.I.T. Press.

Wiener, N. (1950). *The human use of human beings: Cybernetics and society.* Boston, Massachusetts: Avon.

Wolanski, N. (1989). Human life and culture: Dynamic components of ecosystems. *Zygon, 24,* 401-427.

Wu, L. S.-Y. (1976). Some emergent mathematical properties of natural populations. In J.D. White (Ed.), *General systems theorizing: An assessment and prospects for the future,* pp. 211-213. Washington, D.C.: Society for General Systems Research.

©1994 The Institute of Mind and Behavior, Inc.
The Journal of Mind and Behavior
Winter and Spring 1994, Volume 15, Numbers 1 and 2
Pages 35–54
ISSN 0271–0137
ISBN 0-930195-07-8

Deinstitutionalization: The Illusion of Disillusion

Michael McCubbin

Université de Montréal

This article reevaluates the recent tendency to attribute economic causes — cost and fiscal factors — to deinstitutionalization and its subsequent "treatment in the community" mental health system. Economic determinist explanations are shown to be inadequate; instead, the primary impetus behind deinstitutionalization is seen to be the conception of a more humanistic "community care" alternative. How deinstitutionalization was transformed into a mere shadow of that model is explained by analyzing the mediation of social institutions. It is proposed that disillusionment and policy paralysis be replaced with a teleological approach to planning: a long-term strategic plan based on goals and emphasizing the policy environment.

> Disillusion can become itself an illusion
> If we rest in it.
>
> (T.S. Eliot, 1958, *The Cocktail Party*,
> Act 2, p. 138)

Over the last decade the mental health policy literature has unveiled a myriad of political, structural and economic factors that contributed to asylum depopulation and helped to shape the subsequent "treatment in the community" mental health system. The failure to establish the originally envisioned "community care" system in an era of financial restraint has led to a tendency to revise, in hindsight, the perceived impetus behind deinstitutionalization. This impetus is now often seen as economic (cost and fiscal factors), rather than humanitarian (changed attitudes toward the mentally ill and new paradigms placing the problem of mental illness in a broader psy-

---

I am grateful for the helpful comments and encouragement received from David Cohen and Frédéric Lesemann during preparation of this paper. Requests for reprints should be sent to Michael McCubbin, Groupe de recherche sur les aspects sociaux de la santé et de la préventation (GRASP), Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, Canada H3C 3J7.

chosocial context). This article suggests that it would be a misreading of even that literature which emphasizes cost and fiscal arguments to conclude with economic determinist explanations. Instead, I suggest that much of the "economic" impetus for deinstitutionalization was shaped by humanitarian factors, and that the failure to establish community care is not evidence of the weakness of the humanitarian explanation but rather of the inertia of the system due to the power and interests of various social institutions.

Aside from their explanatory weaknesses, economic determinist arguments can create cynicism, disillusion and policy paralysis. Events understood as the consequence of economic "forces" rather than of decisions made by human agents, given their interests and the social structures within which they operate, seem beyond the reach of policy. This article will conclude with the observation that policy efficacy is possible — i.e., that real progress toward a humanitarian, tolerant, psychosocial community care system is attainable — *if* policy directs itself to the structure of the mental health policy system itself. The powers of government bureaucracies, issue advocacy groups, pharmaceutical companies, and the psychiatric profession, mediate and distort well-intentioned incremental reforms. It is time for reformers to pay concerted attention to the nature of the system *itself* as a prerequisite to achieving reform *within* the system.

Background

For at least twenty years the mental health policy literature has focused on deinstitutionalization as the major system "event" in the second half of this century. Reforms begun during the 1960s raised the hopes of many that a new "biopsychosocial" paradigm of "mental illness" and its treatment was on the horizon, ushering in more humane forms of treatment that regarded the patient as a whole person with various psychosocial needs that require attention, not primarily or necessarily to cure the "illness," but to facilitate the patient's normalized participation within society, thereby reducing the harmful impact of symptoms. The paradigm envisaged an array of support services, including counselling, advocacy, housing, transportation, social and home care skills training, education, income, and a variety of therapy alternatives; the core of the system would be the patients' needs or choices, aided by advocacy and means for participation in the system, and integrated for the patient through programs such as case management and community centres. There was broad acceptance of this model of community care, pursuant to an ecological model of mental health, at least in terms of public pronouncements of major mental health system actors (Bloche and Cournos, 1990; Hollingsworth, 1992; see Tyhurst, Chalke, Lawson, McNeel, Roberts et al., 1963).

By the 1980s, at the end of the era of precipitous declines in asylum populations, analysts began displaying a disquietude as to the impacts of the deinstitutionalization movement for patients and society (e.g., Beck and Parry, 1992; Boudreau, 1986; Callahan, 1984; Isaac and Armat, 1990; Rachlin, 1989; Toews and Barnes, 1986). Deinstitutionalization was increasingly perceived as a partial or complete failure: in this regard the attempt at social engineering undertaken in the mental health field suffered from disillusion similar to that following other major social policy initiatives such as the "war on poverty" (Lesemann, 1986).

Some of the major ideological components of the reform movement had become distorted: mental health services, even when labelled "community care," were rather "treatment in the community" (Goodwin, 1989) — where intervention consisted almost entirely of biomedical psychiatric treatment provided to out-patients or short-term general hospital patients (Bachrach, 1981).[1] A tacit assumption of the mental health system was that there were already adequate family supports for outpatients to draw on, or that physically locating them "in the community," with minimal personal entitlements to medical treatment and income, somehow substituted for the wide range of psychosocial and economic supports envisioned under the full community care model.

Today, after a brief flowering of diversity within psychiatry, that profession has become almost exclusively a technical specialty of mainstream biomedicine. The medical model orientation is stronger than ever before in psychiatric practice (Breggin, 1991; Cohen and Cohen, 1986; MacLennan, 1989). As a consequence, the team or case management models of professional intervention under the community care model have become hierarchical structures controlled by psychiatrists or psychiatric conceptions (Hollingsworth, 1992; Prior, 1991; Regan, 1987). Active patient participation has become symbolic or nonexistent.

Economic Explanations of Deinstitutionalization

Recently, there has been a growing tendency to downplay the positive forces behind deinstitutionalization noted by earlier writers including Brown (1985), Foley (1975), and Rochefort (1984): the search for more humane interventions, combined with greater public tolerance for relatively harmless deviances. In parallel, economic forces have been stressed. Correlations between asylum population declines and changes in funding and entitlement programs have suggested to some analysts that "the most powerful and imme-

---

[1] Although data on non-medical support are sparse and dated, this is the case even for clients included in community support programs (Kiesler and Sibulkin, 1987, pp. 196–199).

diate impetus was fiscal" (Bloche and Cournos, 1990, p. 393). The obvious strength of this impetus, combined with other factors such as declines in public spending, the appearance of chronic homelessness of many ex-patients, and the failure to establish community care programs, has created the cynical impression that cold economic facts, rather than the laudable objectives of social reform, were behind the process. For example, Dain (1989) concluded:

> The speed with which the vast system of state mental hospitals, despite all the latter's financial resources and political connections, could be virtually denuded of patients bespeaks more of a desire of state governments to save money and a lack of public support for hospitals than a great faith in alternatives . . . . (p. 7)

However, identification of causes as economic raises the question of why economic incentives or constraints should have changed. The remainder of this section suggests that economic factors often mask more fundamental changes in mental health policy and attitudes.

Cost Arguments

The attitude of cynicism regarding the forces behind deinstitutionalization is noted by Hollingsworth: ". . . many critics of mental health policy suggest that the continuing interest in minimizing mental hospital care is as much an expression of cost control as a concern for patient welfare" (1992, p. 909). Johnson (1990) provides a typical example of this suggestion:

> Depopulation of the state mental hospitals did indeed take place, but not really because more enlightened public attitudes toward the mentally ill made it possible to relocate them to more suitable settings within the community — it took place because the states couldn't afford to provide lifetime care for a huge and growing chronic caseload inside enormous, crumbling hospitals built in the nineteenth century, which proved to be extremely expensive to run at twentieth-century prices. (p. xxii)

While asylum populations were increasing for several decades up to the 1950s (Mechanic and Rochefort, 1990), this fact alone cannot explain subsequent depopulation due to the costs of the system. Indeed, the "cost" argument is explicitly or implicitly based on non-economic factors; e.g., "There have been frequent suggestions that deinstitutionalization was financially inspired, that one state after another turned people out of mental hospitals because it was too costly to provide the level of care *mandated by courts*" (Hollingsworth, 1992, pp. 907–908, italics added).

The influence of legal action in improving asylum conditions, and thereby increasing costs, has been somewhat underrated and forgotten. While only an extremely small number of patients were able to obtain legal help, decisions were precedent-setting and had "class action" impacts. One of the first major impacts during the 1950s was upon hygiene standards. As a result, overcrowded, unsanitary institutions suddenly were faced with the need for major expansion, renovation, or replacement. Subsequently, as a result of further legal and political action, the asylums had to increasingly dispense with the use of physical constraints, requiring more costly alternatives: more staff, fewer patients, redesigned environments. By the early 1960s, "right to treatment" cases were being filed seeking to require asylums to not only maintain residents in adequate physical conditions, but also provide "psychiatric care" (Wald and Friedman, 1978). In retrospect, these legal outcomes have to be regarded as a significant gain for the psychiatric profession, while imposing increased costs on public and private asylums (Bassuk and Gerson, 1978; Brown, 1984).

Insofar as increased costs per patient were a result of increased pressure from advocates, families and the public for both more humane maintenance and psychiatric care, identification of cost as the primary instigating factor for deinstitutionalization would be misleading. Rather than an economic argument, therefore, this becomes a social and political argument based on society's changing expectations concerning appropriate care, and on the increased power of patients and advocates in gaining access to legal and political levers.

The cost–benefit ratio of asylum care did increase during the deinstitutionalization era, due to legal reforms requiring higher standards of care. Alternatives for less severely impaired patients became more attractive, once rudimentary "treatment in the community" approaches were set up. Setting up a new system for a new type of market requires major capital investments: not only in terms of staff and facilities, but also in terms of the intellectual capital expended in conceptualizing and planning the new approaches. It is reasonable to suppose that the timing of the move to asylum alternatives was influenced by "lumpy" capital formation and economies of scale; i.e., the equipment, facilities and staff required for a "treatment-in-community" model (and much more so for the "community care" model), are not infinitely divisible into small fractions. Jones (1988) suggests that a community mental health centre needs to be staffed for service to a population of 100,000 to 200,000. It is likely that factors favouring the creation of a new mental health market came together at a time when a "watershed" was reached at which there were sufficient potential new customers to implement the market. Again, it would be far too narrow a perspective to view this as an asylum cost issue. Rather, the preparation of a society undergoing a paradigm change met the opportunity created by the improving economic feasibility of care in the community.

It is possible that perceptions of costs and benefits of asylum care changed due to changed ability to gather and calculate information (arguably an economic factor), or due to changes in tastes of the public and the decision-makers (a non-economic factor). The former might be manifest in an explanation of deinstitutionalization as the result of science: e.g., the ever more refined tools of health policy analysis finally revealed that it was "inefficient" from a public policy standpoint to retain so large a proportion of those considered mentally ill in institutions. This would provide the only reasonable support, among the cost-related issues discussed above, for emphasizing the cost of asylums as the major causal factor behind deinstitutionalization. However, the more influential critiques of asylum efficacy were not applied policy analysis and evaluation technique but instead fundamental challenges to the paradigms underlining our understanding of mental illness and its place in society.

Fiscal Arguments

The fiscal cause argument implies that an inexorable movement out of the asylums was begun due to the inability of governments to bear the costs of the asylum system (Scull, 1979). It is difficult to see how *capacity* to bear costs can be considered as a major depopulation factor. Deinstitutionalization began during the 1950s and 1960s, when North America was enjoying major economic boom years; this was the period of the most rapid growth in the welfare state. New money was massively injected into new health, welfare and income security programs (Interprovincial Conference, 1980).

While inspection of a particular asylum or sub-national political jurisdiction might reveal a public finance explanation for the nonsustainability of asylums, such fiscal difficulties can be attributed to the circumstances facing specific states or provinces over a limited period of time. In such analysis, historians need to avoid uncritically adopting the fiscal rhetoric of politicians. For example, in the landmark *Wyatt v. Stickney* case [1971] (cited in Wald and Friedman, 1978), the Alabama government pleaded that it could not afford the ordered asylum improvements — pleas the judge did not accept after hearing evidence that Alabama had appropriated that year substantial sums to finance a beauty pageant and a sports hall of fame. A crucial but little known threat to the affordability argument is that civil libertarians and their lawyers made a strategic decision to first attack asylums on the basis of standards rather than on involuntary commitment or broad treatment efficacy considerations, believing that this step would hold early promise in forcing system change (Wald and Friedman, 1978).

In general, during the 1960s and early 1970s sub-national governments discharged or transferred patients in response to incentives, associated with alternatives, rather than in response to affordability considerations. In both Canada and the United States new policies provided new forms of funding —

subsequently cut or severely restrained during the "supply side" era epitomized by the Reagan presidency — for more decentralized forms of care, to individuals (e.g., Medicare), or simply to other institutions (Bloche and Cournos, 1990; Rochefort and Portz, 1993). Indeed, it appears that the most sudden drops in asylum populations are attributable to "transinstitutionalization": transfer of patients to specialized facilities for elderly, children, veterans, criminals, and hospital chronic care beds (Morrissey, Goldman, and Klerman, 1985; Simmons, 1990).

From the perspective of a particular state or province, cost might be the important factor, but from the perspective of an historical analysis of mutations in the North American mental health care system, attention has to be directed to the reasons for the federal government funding changes. In this regard, as discussed above, the capacity of the national governments to maintain funding for asylums does not appear to have been impaired over the period when deinstitutionalization gathered steam. However, since the 1970s, governments have had increased fiscal difficulties due to lower increases in national product growth and tax revenue and increasing costs in various sectors. This reality points to fiscal problems as not a major cause of deinstitutionalization, but as a barrier to subsequent reinstitutionalization — should we decide the experiment was a failure — and as a barrier to realizing or maintaining the community care alternatives originally envisaged.

Mediation of Social Institutions

Society is not made up of homogeneous marbles where planning can be based on the laws of motion of any small set. Rather, individuals associate with each other within complex networks of interacting and overlapping groups, institutions, and rules. Caring and generous people face a variety of incentives and constraints such that their acts can appear selfish, inhumane and alienating, as illustrated by the prisoner's dilemma (Watzlawick, 1976). This complicates the tasks of predicting the course and evaluating the success of a particular mental health intervention.

The Kennedy administration's *Community Mental Health Centers Act* epitomizes the failed or stalled attempt to achieve a revolution in mental health care by direct government intervention. Many problems in implementing community mental health care ideals have been enumerated in the literature (Clark and Dorwart, 1992; Ehrenreich and Ehrenreich, 1971; Nassi, 1978; Regan, 1987). A major limitation of the plan in the United States was that it was up to the various states to take advantage of the cost-sharing proposal: some chose not to adopt the model, and many others watered it down to their own liking (Castel, Castel, and Lovell, 1982; Hastings, 1986). In both cases the intervention, conscious or unconscious, of various social institutions resulted in the failure or distortion of the initiative. The prevailing narrative regarding this failure does not pay sufficient attention to these social structure elements, placing most emphasis on government failure to provide financing for the non-medical psychosocial components of the model (e.g., Boudreau, 1987; Hollingsworth, 1992; Jones, 1988). The implication of this failure is that a lack of will on the part of the public — society in general — sabotaged the plans (Scull, 1990).

While the continued resolve of the general public to spend money on a new and innovative program was clearly insufficient, we must be careful to distinguish between the inherent sentiments of individuals and how those sentiments are influenced, gathered, weighted and transmitted within the political process. At each step of this process barriers and diversions were erected due to the power of groups and institutions whose interests, or perceptions of interests, diverged from that of the idealized homogeneous "voter" (Marmor and Gill, 1989). Following is a partial list of such unforeseen difficulties.

Government bureaucracies. State and province responsibility for mental health care was usually assigned to a health or hospitals department, even though the philosophy underlying community mental health centres (CMHCs) required a new interdepartmental approach in their planning and implementation (which would greatly reduce the health department role and control of resources). These departments and their political representatives were able to frame issues and guide planning in such a way as to turn CMHCs into hospital wards or attached clinics, run by medical practitioners and hospital administrators (Castel et al., 1982; MacLennan, 1989; Regan, 1987).

Pharmaceutical companies. The power of pharmaceutical companies to influence political decisions by lobbying, and psychiatric practice through selective support for medical school programs, conferences and journals, and various gifts to practitioners, is clearly immense. That it has been effective is demonstrable by the amounts spent and by the fact that psychiatry now relies almost exclusively for its practice on psychoactive drugs (Breggin, 1991). Very little research aims to show how promotion directly influences prescriber behaviour; Orlowski and Wateska (1992) studied a sample of physicians receiving expenses-paid trips to resorts for drug symposia and showed that not only did prescription of the promoted drugs rise markedly, but the prescription of the alternate drugs did not fall. Lauzon (1993) found that the major drug companies spend twice as much on promotion than on research. Only within the last few years have some steps been taken by the medical/psychiatric professions to reduce conflicts of interest created by drug company gifts and support of professional/educational activities (Rosner, 1992). However, in recent decades the power of the pharmaceutical industry, particularly as it has evolved into an oligopolistic multinational set of cartels, has more than offset these limited attempts. Wortis and Stone have expressed their concern about this threat to the integrity of the psychiatric profession in an editorial in a 1992 issue of *Biological Psychiatry:* ". . . professional psychiatric societies such as the American Psychiatric Association, the American College of Neuropsychopharmacology, and our own Society of Biological Psychiatry are becoming increasingly dependent on drug-company support" (p. 847).

A number of studies have suggested that drug advertisements have been aimed at extending the indications for prescription of psychoactive drugs into the "problems of daily living" (de Bakey, 1977; Kleinman and Cohen, 1991; Seidenberg, 1971), a result that appears to have been achieved (Kieffer, 1988). The drug companies have been an important factor in turning the vision of community mental health into little more than a market for their products. Deinstitutionalization clearly benefited the drug companies, since the removal of asylum walls contributed to a blurring of the distinctions between those considered severely and mildly mentally ill, making it easier to reach a vast new market of stressed and troubled people.

There is some evidence that the National Institute for Mental Health, the major United States funding agency in that field, and the National Alliance for the Mentally Ill (NAMI), a patient/family lobbying group in the United States numbering more than 100,000 members, are accepting financial support from pharmaceutical companies (e.g., conference sponsorships, drug "scholarships" to NAMI parents) [Breggin, 1991]. The National Mental Health Association (United States) has been criticized for "grey silence" on patient rights and for accepting millions of dollars from a pharmaceutical company for a "public service" campaign on depression (Oaks, 1993).

Issue advocacy groups. The patients, clients, and survivors of the mental health system have never had a major effective lobby group that was not dominated, initially or eventually, by other interests. A number of reasons explain the difficulty of mental health consumers in representing their own interests, not the least of which are the effects of treatment and stigma and the difficulty in obtaining financial resources compared to other interests (Chamberlin, 1990; McCubbin and Cohen, 1994; Mechanic and Rochefort, 1990). The major patient advocate groups are dominated by the patients' parents, whose interests diverge from those of patients (Cohen and McCubbin, 1990; Scull, 1990). While most parents can be considered to be motivated by feelings of compassion and the desire to see their children helped and their suffering reduced, they also wish to avoid blame for their own shortcomings as parents, avoid embarrassment and stigma associated with having mental illness in the family, and reduce tension and conflict in family interactions. While there is nothing wrong with having such desires, they can be expected to influence parent perspectives such that parents should be regarded as representing their own interests rather than those of their children or other family members.

The family lobby groups have been conspicuous for their lack of pressure to implement the psychosocial aspects of the community care model. Instead, they vociferously put forward the idea that mental illness is simply a brain disease (frequently referred to today as "chemical imbalance"), and lobby for funding of biochemical rather than psychosocial research. As is obvious with the gun and anti-abortion lobbies in the United States, relatively small groups with highly committed and well-organized members, and good financing, can have an influence on policy far out of proportion to their membership. Insofar as this power is due simply to intensity of preferences, it can be justifiable in a well-designed political system (Congleton, 1991). However, if such power results from a major structural imbalance in access to the means for democratic participation, the political system is malfunctioning, such that policy inputs like the *Community Mental Health Centers Act* are severely distorted as institutional outputs (Jones, 1988), expanding anomalies in the political bargaining structure (Frey and Eichenberger, 1991). Without super-systemic correction, power begets more power, resulting in an increasingly irrational system.

The psychiatric profession. During the 1960s the psychiatric profession supported the transition from asylum to community mental health care, including the provision of psychosocial assistance. This support was consistent with their interests (MacLennan, 1989). In the years prior to deinstitutionalization psychiatrists were serving two main markets.

One market was the institutionalized population. Treatment of those persons provided relatively little in the way of financial remuneration, professional satisfaction, or status within the community (Scull, 1990). Most asylums were severely underfunded compared to hospitals (Carroll, 1969), the ratio of patient to practitioner extremely high (Wald and Friedman, 1978), and the public perception of asylum psychiatrists turned increasingly negative (Dain, 1989).

The other main market consisted of those well-to-do people and their families who were able to pay for psychotherapy. This had been a solidly growing market, but limited insofar as only the financially capable — or those able to obtain insurance — could access the services (Rochefort and Portz, 1993). Psychotherapy also acquired a tarnish during the 1960s as media portrayals often satirized its practitioners as ineffective, unscientific, and mercenary (Gabbard and Gabbard, 1987). In addition, the field was being invaded by other professions and even by non-professional practitioners.

Community care opened up a huge new market, ostensibly for "social psychiatry," financed by the public purse. Not only could many former asylum inmates be treated in environments more conducive to the psychiatrist's comfort and status, but further expansion of diagnostic labelling could bring in new categories of patients including the stressed, anxious, and well-dressed, most of whom would not have been in the mental health system in prior decades.

In retrospect, it would appear that initially supporting a new model of community care was a strategically wise choice for the psychiatric profession. The vision of how that model might look down the road motivated the government in reducing asylum populations, and the provision of financing for out-patient psychiatric care facilitated access of psychiatrists to patients in the community. Once psychiatrists reached this community market, however, the considerations facing them were changed. Psychiatrists now had the prospect of not only serving a greatly expanded market, but of obtaining the scientific status which accrued to the medical specialties, and protecting their turf, which was being increasingly threatened by other professions active in the fledgling CMHCs (Light, 1985). The psychiatric profession then began to orient toward a biochemical model and sought to replace the idea of "community centre" with that of "clinic in the community," modelled after hospitals or simply run as adjuncts to or wards of hospitals. This orientation has been supported by the tendency of governments to treat mental health policy as a subset of a health policy centered around the hospital which "is regarded as the doctors' workshop — which others pay for but physicians control" (Kiesler, 1992, p. 1080).

Further, the increasing availability and acceptance of psychoactive drugs allowed psychiatrists to function like other medical doctors, dispensing drugs as an act of symbolism enhancing the mystique of the profession (Illich, 1977; Zola, 1978; see Haas and Shaffir, 1982; Lupton, 1993) and providing a mechanism for coping with the uncertainties inherent in the objectives of the profession (Gerrity, Earp, DeVellis, and Light, 1992). Furthermore, the drugs reduced various symptoms considered to be problematic. The ability to actually alter behaviour with drugs allowed for the new narrative of the psychiatric profession to arise: that even if the illnesses were not yet being cured, suffering was reduced and the approach was on the right track to eventually discovering biochemical or genetic markers for which more precise treatments could be devised. This narrative is not only reinforced by families and drug companies, but also by the media, which for many years has trumpeted the latest "discovery" of a schizophrenia gene — while the subsequent replication failures or retractions of earlier findings have received little or no press (Horgan, 1993).

The Further Commodification of Mental Health

Due to the mediation of powerful social institutions such as government bureaucracies, drug companies, families, and the psychiatric profession, another unfortunate consequence of deinstitutionalization was the further commodification of mental health: new products, services and customers were created (Ehrenreich and Ehrenreich, 1971; Ingleby, 1985). That new market came to be served by the "treatment in the community" model. The process of commodification, according to Renaud (1978), invariably follows major social engineering attempts of governments to improve the health systems of capitalist countries. Certainly, the existence of markets in the health field, the entrepreneurial nature of the psychiatric profession, the monopoly powers accorded to the professions, and the exceptional power of medical professions to avoid external regulation — conditions which are generally more marked in the United States than in Canada — impair the ability of governments to "rationally" restructure the mental health system.

The Policy Problem

Collective Paralysis

The major events characterising the rise and stall of the community care model took place over a mere ten-year period, roughly between 1965 and 1975. Pessimistic analyses, often sounding like post-mortems, began immediately after. As the years pass we gain historical perspective: not only do we have the benefit of the experience from intervening years, but as past events become more distant we tend to place them in an earlier as well as later historical context. Hence we can expect the community care initiatives to be subject to continuous or periodic reevaluation in the future.

Too narrow a temporal perspective can result in interpretations which are not only spurious, in the statistical sense, but overly confined to determinisms and fatalisms. A tendency in the recent mental health policy literature sees deinstitutionalization as largely the result of economic and fiscal factors, a view which is reinforced by the failure of governments to spend the money necessary to complete the community mental health care system. The implication of this narrative is that "we can't afford it," that the ideas behind the model have been proven false or unworkable, or that we, as a society, feel that "it's not worth it."

The overly pessimistic prevailing postmortem analysis of the consequences of deinstitutionalization seems to suggest the inevitability of the process that resulted in a technocratic care system that operates "in the community" by simply dispensing drugs (e.g., Johnson, 1990). The feeling seems to be that uncaring governments and the public cloaked cost-saving measures with rhetoric about tolerance and normalization, and about the enabling and caring potential of communities. It is perceived that the public never had the will to complete what it pretended to set out to do.

Even if there is some truth in this perspective, it has created problems for society because it orients us to our failures and selfishness. The analysis of how social institutions such as drug companies, families, and health professionals have adapted to and guided policy interventions in their environments has revealed how relatively narrow self-interest can sabotage reforms whose objectives are shared by a large number of individuals or higher-level collectivities. Indeed, aside from the diversity of opinion within groups — a diversity which groups generally try to suppress — an individual may have many conflicting identities: e.g., the psychiatrist who feels, given the options currently available, that drug treatment is preferred, that his or her profession should lobby for more money to be directed to genetic research into schizophrenia, yet at the same time, both as a psychiatrist and as a person, feels that the mental health system would be better if it adopted a much more ecological approach to etiology, research, treatment, care, and support. The problem is a public choice problem, characterized by the paradox of the prisoner's dilemma. Individuals chart short-term, self-interested strategies, eschewing the strategies they might plot as members of higher-order collectivities, perceiving that they have little influence on the "system," and that other individuals will also act in the same way.

For a brief period during the 1960s, various elements coalesced and increased optimism that, as a society, we could successfully work together to achieve higher level collective goals. It was in this period of collective empowerment that both the rhetoric and experiments of community care had their heyday, inspired generally by social thought and movements envisaging a society of greater tolerance, wider participation, and a more ecological concept of the world and humanity's place in it, and by counterparts of these social developments in critiques of the asylum system as dehumanizing. The difficulties encountered (often resulting in worsened circumstances for those labelled mentally ill, especially those more severely impaired or marginalized, who found themselves homeless, suffering from iatrogenic diseases, or still warehoused in institutions, whether or not they are called asylums) have reduced our confidence that we can solve social problems and act collectively, as communities, to progress toward goals — explicit or amorphous — which we feel in our heart(s) are more just, humane, and ethical, as well as efficacious over the long-run.

Major reform of the mental health system does indeed seem virtually hopeless today, unless individually and collectively we begin to accept responsibility for the system as a whole. This means not only acknowledging its failures as our own, but also accepting some credit for our tentative initial steps toward reform. This would mean, as proposed by Scull in referring to the "lunacy reform movement" of over a century ago, accepting that there has been an "authentic shift in moral consciousness," and rejecting the "crude reductionism" which seeks to explain the humanitarian sensibility as produced by material or economic interests (1985, p. 134).

Toward a Teleological Policy Planning Approach

The bulk of this paper has been devoted to an attack on economic determinist arguments — cost and fiscal factors — as the "cause" of deinstitutionalization, and to how social institutions transformed the community care objectives of health planners. Implicit or explicit throughout is a competing argument as to what broad factor is crucial in finding meaning behind the changes in the mental health system: that of actor choice — whether actors be conceived as individuals or collectivities. The emphasis on choice implies a teleological rather than determinist perspective, if choice is understood, as it is here, as being freely made within generally non-binding constraints.

Incremental policy. Simmons (1990) has described the development and implementation of mental health policy in Ontario as "incremental," noting that the future success of community mental health care "is by no means assured" (p. 267). Particularly because of the rise in recent years of a number of powerful non-client constituencies, Simmons argues the necessity of long-range strategic government planning. The mental health policy system has become increasingly unpredictable, particularly as a result of incremental changes, each of which not only modifies service institutions and delivery mechanisms, but creates a new political environment. Hence the system suffers from the "Butterfly Effect" — as the term has been used by Gleick (1987) to illustrate the difficulties in weather and macroeconomic forecasting and planning.

It is becoming increasingly clear that not only do client interests suffer during the minor policy skirmishes, but that the compromises resulting from incremental policy making can greatly alter long-term trajectories. For example, locating community care clinics in hospitals as an "interim" step has created new vested interests that may render further transformations increasingly difficult to achieve. There is no good reason, theoretically or given our experience in social engineering, to think that an incremental approach provides "steps in the right direction." Every change in the system creates a new environment; changed social structures change constraints which change preferences (Etzioni, 1988). Furthermore, social institutions seek to maintain their identities in adapting to interventions by cooptation (Ehrenreich and

Ehrenreich, 1978) and other compensating adjustments which, according to Crozier and Friedberg (1977), more or less totally transform the sense of reforms.

Political strategy. Therefore, a fundamental element of a strategic plan is a *political* strategy. As opposed to technocratic planning which merely inserts new structures within existing social and economic contexts, a political strategy would recognize the power and interests of the various actors within the system with a view to introducing or enhancing constraints or incentives to facilitate policy implementation. At this stage it appears that the emphasis should not be on mental health policy in a narrow sense but on the policy *environment*. While the purpose of this paper is not to develop policy proposals but to help clarify our understanding of the circumstances within which policy is made and distorted, this paper raises certain issues which should be addressed by a strategic policy emphasizing an improvement of the policy environment. Are there ways to reduce the stigma of ascribed mental illness, felt by clients and families, without overemphasizing biochemical attributions? Can the financial power of pharmaceutical companies to promote their products and influence policy and practice be justified? Are academics and medical practitioners, universities and professional bodies, excessively vulnerable to at least the appearance of conflicts of interest as a result of pharmaceutical funding? Can the formation of mental health policy within governments escape the umbrella of *health* (particularly insofar as health policy is concerned with hospitals and medicine)? How can goverments integrate for distressed clients a wide variety of social services, when those services are administered under separate bureaucracies?

The Will to Change

As Bateson (1972) suggested, the most intractable problems may require for solution a leap of faith, an almost paradoxical effort to emerge into a higher level of consciousness — in effect, to learn how to change our own identity, becoming dynamic rather than static personalities. This advice is applicable to the mental health policy field: we need to retain sight of our goals, the ethical outcomes we seek, and avoid demoralization due to dwelling excessively upon static and deterministic constructions of our present and past.

There is an obvious irony in this article's proposal to address the policy environment in order to reduce the influence of powerful interests which have already shown their ability to thwart major policy initiatives in the past. Given the disillusionment resulting from our inability thus far to achieve the objectives of community care, a revitalization of our collective will is called for by reopening the most fundamental questions that must underly policy for its expression to be coherent. Are we really more willing to tolerate the deviances that we label "mental illness"? Do we yet recognize distressed persons as multi-dimensional social beings with needs that can only be met by and within community? Is the "community care" approach simply another service modality designed to meet a broader variety of needs, or does it reflect a more pragmatic view of "mental illness" as human problems?

Refocusing on these questions is an essential prerequisite to achieving true mental health policy reform. However, even if society were able to commit itself to a radically different paradigm of "mental illness," the actual power arrangements of the system are bound to divert us away from our ultimate goals. This article strongly suggests, therefore, that the mental health policy literature must progress beyond detailing the litany of incremental policy failures, to explicit consideration of how the distribution of power constraining the system can and should be altered. Those with power within the system must not, *by default*, be allowed to shape the mental health policy of a fair and democratic society.

References

Bachrach, L.L. (1981). General hospital psychiatry: Overview from a sociological perspective. *American Journal of Psychiatry*, *138*, 879–887.

Bakey, L. de (1977). Happiness is only a pill away: Madison Avenue rhetoric without reason. *Addictive Diseases*, *3*, 273–286.

Bassuk, E.L., and Gerson, S. (1978). Deinstitutionalization and mental health services. *Scientific American*, *138*, 46–63.

Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballantyne Books.

Beck, J.C., and Parry, J.W. (1992). Incompetence, treatment refusal, and hospitalization. *Bulletin of the American Academy of Psychiatry and Law*, *20*, 261–267.

Bloche, M.G., and Cournos, F. (1990). Mental health policy for the 1990s: Tinkering in the interstices. *Journal of Health Politics, Policy and Law*, *15*, 387–411.

Boudreau, F. (1986, March). From asylum to mental health: Is the patient any better off? *Canada's Mental Health*, pp. 16–18.

Boudreau, F. (1987). The making of mental health policy: The 1980s and the challenge of sanity in Quebec and Ontario. *Canadian Journal of Community Mental Health*, *6*, 27–47.

Breggin, P.R. (1991). *Toxic psychiatry*. New York: St. Martin's Press.

Brown, P. (1984). The right to refuse treatment and the movement for mental health reform. *Journal of Health Politics, Policy and Law*, *9*, 291–313.

Brown, P. (1985). *The transfer of care: Psychiatric deinstitutionalization and its aftermath*. Boston: Routledge & Kegan Paul.

Callahan, J.C. (1984). Liberty, beneficence, and involuntary confinement. *Journal of Medicine and Philosophy*, 9, 261–293.

Carroll, H.A. (1969). *Mental hygiene: The dynamics of adjustment*. Englewood Cliffs, New Jersey: Prentice–Hall.

Castel, R., Castel, F., and Lovell, A. (1982). *The psychiatric society*. New York: Columbia University Press.

Chamberlin, J. (1990). The ex-patients' movement: Where we've been and where we're going. *Journal of Mind and Behavior*, *11*, 323–336.

Clark, R. E., and Dorwart, R.A. (1992). Competition and community mental health agencies. *Journal of Health Politics, Policy and Law*, *17*, 517–540.

Cohen, D., and Cohen, H. (1986). Biological theories, drug treatments, and schizophrenia: A critical assessment. *Journal of Mind and Behavior*, *7*, 11–36.

Cohen, D., and McCubbin, M. (1990). The political economy of tardive dyskinesia: Asymmetries in power and responsibility. *Journal of Mind and Behavior*, *11*, 465–488.

Congleton, R.D. (1991). Information, special interests, and single-issue voting. *Public Choice*, *69*, 39–49.

Crozier, M., and Friedberg, E. (1977). *L'acteur et le système*. Paris: Éditions du Seuil.

Dain, N. (1989). Critics and dissenters: Reflections on "anti-psychiatry" in the United States. *Journal of the History of the Behavioral Sciences*, *25*, 3–25.

Ehrenreich, B., and Ehrenreich, J. (1971). *The American health empire: Power, profits and politics*. New York: Vintage Books.

Ehrenreich, B., and Ehrenreich, J. (1978). Medicine and social control. In J. Ehrenreich (Ed.), *The cultural crisis of modern medicine* (pp. 39–79). New York: Monthly Review Press.

Eliot, T. S. (1958). *The cocktail party*. London: Faber and Faber.

Etzioni, A. (1988). *The moral dimension: Toward a new economics*. New York: The Free Press.

Foley, H.A. (1975). *Community mental health legislation: The formative process*. Toronto: Lexington Books.

Frey, B.S., and Eichenberger, R. (1991). Anomalies in political economy. *Public Choice*, *68*, 71–89.

Gabbard, K., and Gabbard, G.O. (1987). *Psychiatry and the cinema*. Chicago: University of Chicago Press.

Gerrity, M. S., Earp, J.A.L., DeVellis, R.F., and Light, D.W. (1992). Uncertainty and professional work: Perceptions of physicians in clinical practice. *American Journal of Sociology*, *97*, 1022–1051.

Gleick, J. (1987). *Chaos: Making a new science*. New York: Penguin Books.

Goodwin, S. (1989). Community care for the mentally ill in England and Wales: Myths, assumptions and reality. *Journal of Social Policy*, *18*, 27–52.

Haas, J., and Shaffir, W. (1982). Taking on the role of doctor: A dramaturgic analysis of professionalization. *Symbolic Interaction*, *5*, 187–203.

Hastings, M.M. (1986). The long-term care puzzle and mental health policy. *American Behavioral Scientist*, *30*, 143–173.

Hollingsworth, E.J. (1992). Falling through the cracks: Care of the chronically mentally ill in the United States, Germany, and the United Kingdom. *Journal of Health Politics, Policy and Law*, *17*, 899–928.

Horgan, J. (1993). Eugenics revisited: Trends in behavioral genetics. *Scientific American*, *268*, 122–128.

Illich, I. (1977). *Limits to medicine: Medical nemesis. The expropriation of health*. New York: Penguin Books.

Ingleby, D. (1985). Mental health and social order. In S. Cohen and A. Scull (Eds.), *Social control and the state* (pp. 141–188). Oxford: Basil Blackwell.

Interprovincial Conference of Ministers Responsible for Social Services. (1980). *The Income Security System in Canada*. Ottawa: Canadian Intergovernmental Conference Secretariat.

Isaac, R.J., and Armat, V.C. (1990). *Madness in the streets: How psychiatry and the law abandoned the mentally ill*. New York: Free Press.

Johnson, A.B. (1990). *Out of bedlam: The truth about deinstitutionalization*. New York: Basic Books.

Jones, K. (1988). *Experience in mental health: Community care and social policy*. London: Sage Publications.

Kieffer, G.H. (1988). The control of behavior. In R.B. Edwards and G.C. Graber (Eds.), *Bioethics* (pp. 352–374). Orlando, Florida.: Harcourt Brace Jovanovich.

Kiesler, C.A. (1992). U.S. mental health policy: Doomed to fail. *American Psychologist*, *478*, 1077–1082.

Kiesler, C.A., and Sibulkin, A.E. (1987). *Mental hospitalization: Myths and facts about a national crisis*. London: Sage.

Kleinman, D.L., and Cohen, L.J. (1991). The decontextualization of mental illness: The portrayal of work in psychiatric drug advertisements. *Social Science and Medicine*, *32*, 867–874.

Lauzon, L.-P. (1993, 21 July). Six compagnies pharmaceutiques ont réalisé des profits de 22,3 milliards $ en trois ans. *Le Devoir*, A5.

Lesemann, F. (1986). La pauvreté: raison d'État, affaire de coeur. *Revue internationale d'action communautaire*, *16/56*, 3–5.

Light, D. (1985). Professional training and the future of psychiatry. In P. Brown (Ed.), *Mental health care and social policy* (pp. 218–236). Boston: Routledge & Kegan Paul.

Lupton, D. (1993). The construction of patienthood in medical advertising. *International Journal of Health Services*, *23*, 805–819.

MacLennan, D. (1989). Psychiatric challenges to the asylum: A theme in the development of Canadian psychiatry, 1918–1963. *Canadian Journal of Community Mental Health*, *8*, 75–91.

Marmor, T., and Gill, K.C. (1989). The political and economic context of mental health care in the United States. *Journal of Health Politics, Policy and Law*, *14*, 459–475.

McCubbin, M., and Cohen, D. (1994, June). *A political economy of the mental health system: Power disparities and interest divergences between clients and psychiatry*. Paper presented to the xxth International Congress of Law and Mental Health, Montreal.

Mechanic, D., and Rochefort, D.A. (1990). Deinstitutionalization: An appraisal of reform. *Annual Review of Sociology*, *16*, 301–327.

Morrissey, J.P., Goldman, H.H., and Klerman, L.V. (1985). Cycles of institutional reform. In P. Brown (Ed.), *Mental health care and social policy* (pp. 70–98). Boston: Routledge & Kegan Paul.

Nassi, A.J. (1978). Community control or control of the community? The case of the community mental health center. *Journal of Community Psychology*, 6, 3–15.

Oaks, D. (1993). (Ed.). Clue: Government & industry & phoney "consumer" groups unite. *Dendron*, *33/34*, 20.

Orlowski, J.P., and Wateska, L. (1992). The effects of pharmaceutical firm enticements on physician prescribing patterns. *Chest*, *102*, 270–273.

Prior, L. (1991). Community versus hospital care: The crisis in psychiatric provision. *Social Science and Medicine*, *32*, 483–489.

Rachlin, S. (1989). Rethinking the right to refuse treatment. *Psychiatric Annals*, *19*, 213–222.

Regan, T.G. (1987). Negotiating an innovative mental health service: The question of limits to negotiated orders. In D. Coburn, C. D'Arcy, G.M. Torrance, and P. New (Eds.), *Health and Canadian society: Sociological perspectives* (pp. 501–514). Markham, Ontario: Fitzhenry & Whiteside.

Renaud, M. (1978). On the structural constraints to state intervention in health. In J. Ehrenreich (Ed.), *The cultural crisis of modern medicine* (pp. 101–120). New York: Monthly Review Press.

Rochefort, D. A. (1984). Origins of the "third psychiatric revolution": The Community Mental Health Centers Act of 1963. *Journal of Health Politics, Policy and Law*, 9, 1–30.

Rochefort, D.A., and Portz, J.H. (1993, Spring). Different systems, shared challenges: Assessing Canadian mental health care from a U.S. perspective. *American Review of Canadian Studies*, pp. 65–82.

Rosner, F. (1992). Ethical relationships between drug companies and the medical profession. *Chest*, *102*, 266–269.

Scull, A. (1979). *Museums of madness: The social organisation of insanity in nineteenth century England*. London: Allen Lane.

Scull, A. (1985). Humanitarianism or control? Some observations on the historiography of Anglo–American psychiatry. In S. Cohen and A. Scull (Eds.), *Social control and the state* (pp. 118–140). Oxford: Basil Blackwell.

Scull, A. (1990). Deinstitutionalization: Cycles of despair. *Journal of Mind and Behavior*, *11*, 301–312.

Seidenberg, R. (1971). Drug advertising and the perception of mental illness. *Mental Hygiene*, *55*, 21–31.

Simmons, H. (1990). *Unbalanced: Mental health policy in Ontario, 1930–1989*. Toronto: Wall & Thompson.

Toews, J., and Barnes, G. (1986, June). The chronic mental patient and community psychiatry: A system in trouble. *Canada's Mental Health*, pp. 2–7.

Tyhurst, J.S., Chalke, F.C.R., Lawson, F.S., McNeel, B.H., Roberts, C.A., Taylor, G.C., Weil, R.J., and Griffin, J.D. (1963). *More for the mind: A study of psychiatric services in Canada*. Toronto: Canadian Mental Health Association.

Wald, P.M., and Friedman, P.R. (1978). The politics of mental health advocacy in the United States. *International Journal of Law and Psychiatry*, *1*, 137–152.

Watzlawick, P. (1976). *How real is real? Confusion — disinformation — communication*. New York: Random House.

Wortis, J., and Stone, A. (1992). Editorial: The addiction to drug companies. *Biological Psychiatry*, *32*, 847–849.

Zola, I.K. (1978). Medicine as an institution of social control. In J. Ehrenreich (Ed.), *The cultural crisis of modern medicine* (pp. 80–100). New York: Monthly Review Press.

Mental Health Policy from a Systemic Perspective

Orientations for Strategic Reform

Short title: GST & Mental Health Reform

Michael McCubbin* and David Cohen

Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), Université de Montréal

* Corresponding author.
GRASP, Université de Montréal,
C.P. 6128, succursale Centre-ville,
Montréal (Québec), Canada H3C 3J7
[redacted]

Word count: abstract 207, text 9916, footnotes 195, references 2908

**Submitted to *Journal of Health Politics, Policy and Law*,
25 November 1997**

Abstract—Most writers now recognize that mental health policy and the mental health system are extremely resistant to real changes that reflect genuine biopsychosocial paradigms of mental disorder. Writers bemoaning the intransigence of the mental health system tend to focus on a small analytical level, only to find themselves mired in the rationalities of the existing system. Problems are acknowledged to be system-wide, yet few writers have used a method of analysis appropriate for systemic problems.

This article describes how General Systems Theory (GST) has thus far influenced the study of the mental health policy and politics system, and argues that a GST perspective is profitable for reconceiving the mental health system, enabling a fresh basis for the development of reform strategies. The mental health system should be seen as a social system influenced by larger political and economic dimensions, not just as a "delivery system" scientifically constructed by neutral experts.

Furthermore, the policy planning process should be viewed as part and parcel of a mental health system modeled as complex and dynamic. The systemic perspective outlined here should help both to clarify the value-based objectives that we hold for the system and, consequently, to plan for the strategic reforms that have so far eluded us.

Keywords—mental health, policy, politics, power, General Systems Theory

Acknowledgments—We would especially like to thank Jocelyne Boivin-Ostiguy, Line Garand, and Andrée Demers, for helping to provide a friendly and effective working environment at GRASP, as well as various needed resources, during the long period of time this article was being developed. An earlier version of this paper was presented to Conférences du Laboratoire du Département d'Anthropologie, Université Laval, Québec, March 1997.

Mental Health Policy from a Systemic Perspective

Orientations for Strategic Reform

Introduction

Drawing upon the General Systems Theory (GST) analytical perspective, this article advances a systematic approach to understand the mental health system and to facilitate the development of reform strategies that recognize the system's complexity and changing nature. The article first discusses the failure of major reform efforts in the mental health system and the limitations of mainstream analysis of mental health politics and policies with respect to the objectives of analysis and reform.

We then argue, drawing upon the sparse literature that has explicitly applied a GST perspective to mental health policy and politics, that a fresh and fruitful basis for planning fundamental reform of the system requires a much broader view of the mental health sector as a system. Such a view respects its complex and dynamic nature, and recognizes policy planning as part and parcel of the system. Furthermore, this view accounts for the political, economic and cultural dimensions which interact with the mental health system.

Mental Health Policy Reforms: Spinning in the Mud

In many western countries the mental health sector has been viewed as a "system in crisis" (Goldie and Fredén 1991; Kemp 1991; Prior 1991)—a situation perpetuated by a series of failures in major reform initiatives (Bloche and Cournos 1990; Kiesler 1992; Marmor and Gill 1989; Mechanic and Rochefort 1990). "Deinstitutionalization" loosely captures the major reform themes over the past 40 years. The ostensible theoretical basis for this process has been a "biopsychosocial" perspective of psychiatric patients as whole persons whose various needs should be met by and within community (Rochefort 1984; Tyhurst et al. 1963). However, despite the apparent initial support of major system actors for a network of community-based resources intended to reduce reliance upon medical and institutional care, attempts to do so have been derailed (Becker 1993; Goodwin 1989; Hollingsworth 1992, 1996; McCubbin 1994a; White and Mercier 1991).

Program frameworks were set up, aiming to ensure a continuum of care and various psychosocial services in communities. However, funding and implementation of these programs has been very slow given initial expectations, with most funding going to inpatient care (Dorvil 1997; Mechanic 1996). Worse, there are recent indications of backsliding toward a new "dark age" of mental health (Cohen C. I. 1997; Thompson 1994). Cutbacks to social spending reduce psychosocial programs to empty shells, whose real function is hospital referral and coercion or monitoring of outpatients to comply with medication regimen. Housing options to severely distressed persons might today include supported or group housing, or rooming house beds that are more or less "supervised", but it is uncommon for patients to have care options as a supplement or alternative to prolonged drug treatment.

Indeed, it is arguable that what began as a humanitarian movement away from asylums toward community support under a biopsychosocial perspective has transformed into a major <u>expansion</u> of a coercive medical care system built around hospitals (Durham and La Fond 1996). Every indicator points to incessant growth in the mental health empire, particularly in the United States, with the targeting of new client categories, including children and those stressed by contemporary upheavals in the workplace (Diller 1996; Kleinman and Cohen 1991). In the U.S., the number of inpatient episodes per 100,000 population has grown from 795 to 917 between 1955 and 1990 (anon. 1994). Rather than integrating emotionally distressed persons within ordinary social environments by providing or facilitating access to community based resources, the system is rather geared to crisis intervention with chronic patients, or to prescribing psychoactive drugs to stressed and troubled people not living in hospitals.

According to popular discourse since the initiation of the major reform attempts, social values driving them were the related concepts of dignity, normalization, empowerment, and human rights (Brown 1984; Wald and Friedman 1978). To some observers, new legal rights, progressive on paper, have meant little for persons faced with forced confinement or treatment (Appelbaum

1997; Holstein 1993); to others, these rights are tying the hands of psychiatrists and impeding the provision of needed care (Isaac and Armat, 1991; Torrey, 1997). It remains that much treatment is implicitly coercive rather than ordered by a court or tribunal, e.g. where its acceptance is a precondition to receive income or other services (Fineman 1991; Wells 1997), or in the common case of "voluntarily" accepting treatment that the patient believes would be otherwise forced (Lurigio and Lewis 1989; Miedema 1994; Reed and Lewis 1990).

New ways have been created to restrain the liberties of persons diagnosed as mentally ill, including "compulsory community treatment orders" (Boudreau and Lambert 1993a and 1993b; Szasz 1994). Yet users themselves, in the rare occasions when asked about what is important to them (open-ended inquiry going far deeper than the usual surveys of "user satisfaction"; see Fitzpatrick and White 1997), highly value autonomy, privacy and choice. The importance of these values to users helps to explain their massive preference to live in the community (Upshur et al. 1997), even though they often describe their lives in community as isolated and bleak (Davidson et al. 1996).

In sum, "deinstitutionalization" has lost most of its allure. Many individuals needing care or psychosocial supports—beyond drugs or a hospital bed—receive them neither in institutions nor in the community. While in this century the numbers of those labeled mentally ill have risen steadily, the prognosis for the more serious troubles and conditions has hardly changed (Hegarty et al. 1994), and there are more than a few indications that "treatment" for them can hurt rather than help: critics have decried the infantilization of patients (Ingleby 1985), the exacerbation of patient trauma (Jennings 1994; Mack 1994), and the magnitude of treatment-induced illnesses (Breggin 1997; Cohen D. 1997).

The State of Policy Analysis in the Mental Health Field

If the scenario described above is accurate, it is perplexing that the mental health policy and politics literature has not mobilized to create a sort of "Manhattan Project" to get reform back on track toward the objectives which have been reinforced in discourse and public policy initiatives at the same time as reality has been going in the opposite direction. The literature has frequently described major nationwide policy failures as well as perverse effects of reform attempts at regional and local levels. Rarely has a serious effort been made to advance new strategic policy orientations that address the scope of the problem as described above. Insofar as it goes beyond description to prescription, policy development still amounts to "tinkering in the interstices"—a practice recognized by its own literature as futile.

The literature tends to be multidisciplinary rather than interdisciplinary. However, policy analysis and development requires several skills that can best be developed in an interdisciplinary context (Shadish Jr. et al. 1989). The literature of the disciplines is hampered by theoretical perspectives and methodologies inapplicable to the policy level. For example, none of the empirically based articles published during 1989-1991 in the Canadian Journal of Community Mental Health used a methodology directed to the community or societal levels, even though a third interpreted at these levels; most articles used methodology directed to the individual level. This suggests that at least in that journal there is a mismatch between methodology and interpretation which, compared with earlier years, has been growing (Peirson and Walsh-Bowers 1993). When problems are too narrowly defined, solutions may be inappropriate, have unforeseen perverse effects, or be impossible to sustain in the long run (Blum 1983; McCubbin 1997).

Those writing about mental health policy should pay more conscious attention to their own method in approaching the phenomena that interest them, and enrich their method with developments in the broader policy sciences. As observed by Rochefort:

> Of the prodigious amount of policy analytic studies in existence, especially those studies appearing in mainstream professional journals, only a relatively small proportion concern themselves with mental health policymaking. Fewer still relate mental health policy development to theoretical issues in the policy research field. The oversight is lamentable, for mental health policy supplies an unusually rich substantive context, apposite both for the refinement and supplementation of current policy concepts and hypotheses (1993: 13).

In general the mental health policy literature is descriptive rather than theory driven, built upon historical case studies (e.g., Foley 1975; Goodwin 1989; Durham and La Fond 1996). The best of this literature is usually chock-a-block with facts, statistics, views of other analysts, narratives, and vignettes. For example, Benson has, without explicit theory, "examined U.S. mental health policy from a broad historical, social, political, and economic perspective ..." (1994: 132). Such studies are useful to construct certain types of arguments, especially along the lines of political economy, but many of them are hampered by their dependence upon "found" data: major events, laws, available statistics, the discourse of those who have visible power or socially authorized influence—politicians, administrators, professionals, corporations, academics, and family lobby groups—but rarely the discourse of patients.

The stringing together of such found data in historical sequence may provide a seemingly coherent narrative which is nevertheless weak in its explanatory ability insofar as fundamental but less visible processes are excluded. Events and structures are explained, in hindsight, by tracing a path backwards in time from that which is explained. This does not establish a chain of causality but a chain of events which appear to be causal (see Macy and Flache 1995). This method has an inherent tendency to attribute the most explanatory power to the most temporally proximate "major" events. One may question whether the types of explanation that have been developed can be applied with any reliability not only to prediction in the usual sense, but to understanding where we are now (which can only be properly conceived of in the context of processes or directions over time). Do we have any good reason to think that what were important factors in the past will be so in the present and future? This is particularly important since society is changing its basic structures, and perhaps attitudes/values, at an accelerating pace. The case studies are replete with interesting observations that need to be placed within a larger theoretical framework explaining how and especially why the system operates as it does.

While mainstream literature frequently discusses the policy process and the inputs of the principal actors and "stakeholders" or "partners", it is not really addressed to the broad questions "What are the politics of mental health?" or "What is the nature of the political environment of the mental health sector?" but rather "How can we explain the existence of 'x' (historical event, structure, policy, law, etc.)?" Using the language of political science, we might describe the literature as centered around formal institutions, even if it is no longer focused solely on them.

The social sciences have largely accepted the notion that "social problems" are social constructions, shaped by culture, politics, economics, "stance," etc. (see Rochefort and Cobb 1994). Many analysts have therefore shifted their attention to the politics and other social processes of problem definition and resolution, aiming to be, in effect, dispassionate observers of ideology (Mannheim 1936). However, one should not lose sight of the possibility that some of what we see as problems may be symptoms, and that dealing only with symptoms does not solve the problems. Too often only "important" or visible actors, processes and goals/values are recognized. While diffuse factors like "culture" are given the obligatory nod, writers attend to the most concrete factors which might, at most, be triggers, or necessary but not sufficient, and which, in recognition of "complexity", are described as "powerful" or "immediate" forces. As a result, obvious blind spots remain, such as collective human agency and underlying fundamental societal structures. These are addressed by rarely cited, often Marxian influenced writers, e.g. Ehrenreich and Ehrenreich (1971), Hill (1983), Hopton (1997) and Kenig (1992). Indeed it is ironic that "radical" neo-Marxist writers appear to be less economic determinist than the "moderate" policy science writers who grasp at economic "forces" for lack of a theoretical paradigm.

Understanding of the mental health system still relies upon the vapours apparently emanating from the objects of study. This is especially so where no explicit theory guides inquiry. In the mental health field the paucity of integrative, systematic theory-driven critical inquiry has resulted in large gaps and distortions in knowledge. In a policy environment where:

- the pharmaceutical industry in the U.S. spent $5 billion a year on drug promotion and $7 billion a year on drug research (Wortis and Stone 1992)[1];
- well-funded family lobby groups like the National Association for the Mentally Ill (NAMI) and Children and Adults with Attention Deficit Disorders (CHADD) are successful in encouraging the funding of largely biological research by public granting agencies like the National Institute of Mental Health (McLean 1990; Mechanic 1996; Merrow, 1995);

---

[1] There are few published analyses of pharmaceutical company spending, and they are hampered by company spending data which fail to clearly separate promotional spending from research and administrative spending. There are indications, based on conservative estimates, that in many countries promotional spending equals or exceeds research spending (Haaijer-Ruskamp and Dukes 1991; Lauzon 1993).

- governments reduce funding of universities and research funding bodies in favour of increased academic "partnership" with the private, para-public, and public sectors;

a vast imbalance in research emphasis is created, favouring medical as opposed to psychosocial approaches, favouring the interests of professions, corporations and families rather than those of clients and survivors. As a result of this imbalance, we do not know enough about how community programs do and should work; insufficient research attempts to develop and apply psychosocial models of care based on fundamental objectives; the "needs" of clients are assumed rather than made the object of careful inquiry (Davidson et al. 1996)—indeed, we have virtually no idea what the "mentally ill" experience, perceive and feel in their daily lives, either as clients of various interventions or as "falling through the cracks".

The widespread dissatisfaction with the mental health system and its long resistance to meaningful reform would seem to require thorough and deep reflection as to

- social objectives for the system,
- contextual and internal forces hindering achievement of the objectives, and
- strategies large enough to break the system's inertia and reshape it in the public interest.

Yet, with rare exceptions the overall system dynamics have not been directly addressed; the focus has been on policy initiatives rather than on the context within which policy is developed, and on countless explanatory variables that do not fit together coherently to allow for explanation rather than description. Despite the frequently acknowledged failures of the system and of attempts to change it, actions are still justified with reference to the intervention standards and good intentions of the helping professions, planners, and administrators, rather than according to results shown to meet broader societal interests and values.

GST Perspectives of the Mental Health Policy and Politics System

Von Bertalanffy (1968) referred to GST as a meta-theory akin to mathematics. As such it has no empirical content. Rather, it provides an epistemological perspective, sensitizing the analyst to questions of structure and interaction. GST provides an approach to consciously construct an abstracted reality, simplifying it while at the same time capturing its multi-dimensionality. Hence, users of modern systems theory in the social sciences tend to highlight issues of complexity (multiple causality), dynamism (the importance of processes within the system and of the mutating nature of the system), uncertainty (actor choices are strategic due to the unknown future and the lack of transparency regarding the intentions of others), and, reflecting the preceding, holism (social systems are viewed as "open", being parts of larger systems)[2].

In this section, we critically review the few extant applications of GST in the mental health policy and politics literature, aiming to show some of the implications of systems thinking for the study of mental health policy and politics. We focus first on GST perspectives in mental health policy processes, then move on to a higher systemic level of analysis, that of the dynamics of mental health system change. We argue throughout that effective policy analysis and development in the mental health field requires more attention to system dynamics than has thus far been the case.

Mental Health Policy Processes

Although implicit or explicit use of systems theory has been common in the public policy and administration literature since the 1960s, and some of its language and concepts have been prominent in social science thinking, it has rarely been applied methodically to the larger mental health policy questions. However, there exist several applications of GST to more localized policy issues dealing with mental

---

[2] An introduction to GST, with some health policy applications, might usefully include: Ackoff 1974; Bahg 1990; Bateson 1972; von Bertalanffy 1981; Bowler 1981; Buckley 1967; Campbell 1982; Capra 1990; Evans and Stoddart 1990; Fontaine 1997; Lane and Jackson 1995; McCubbin 1994b; Morin 1990; Skyttner 1996; Sutherland 1973.

health service delivery and administration (e.g., Austin 1993; Juba 1997; Scott and Black 1986; Ziegenfuss Jr. 1983). At the level of analysis of how policies are forged, the GST approach leads the researcher to consider interactions among subsystems that may not be obvious (or recognized in the literature), and to place them in the contexts of suprasystems surpassing the boundaries of the "mental health sector" as usually understood.

Foley and Sharfstein recognized the urgency to place the elements of the system in the context of larger systems. Commenting upon the 1978 report of the U.S. President's Commission on Mental Health, these authors noted that:

> the report took on a particular systems cast. In this it was reflecting the new thinking in the field, which sought ways to transform the de facto mental health service system into an intentional system that would articulate with other systems of health and human services where these other systems impinged on the treatment and care of the mentally ill ... (1983: 116)

Unfortunately, this "new thinking" seems to have reached its apogee in the mid 1980s, when some members of the new "institutional theory" school (see Scott 1987) too briefly turned their attention to the mental health system. Meyer (1985) made the interesting point that old-style organization theory, which stresses technical bases of organization, has little to offer our understanding the mental health system, since "It is not obvious that there are technical considerations in this field of any consensual status ...". Meyer therefore looked at the "institutional rules that define the bases of organizational rationalization in mental health systems" (1985: 587). Meyer's object is not the mental health system, but rather certain of its abstract subsets such as organizations. These are, however, seen institutionally, i.e., as very open systems largely constituted, legitimated and maintained by the environment.

Meyer did not provide causal arguments per se but an explanation of how the mental health system can maintain its legitimacy in the eyes of society. He provided a brief "problematics of the social construction of mental health" under the headings of goals, technology, resources, and sovereignty. Drawing on this framework, he characterized the American system in comparison to abstract or unstated other systems (using criteria such as numbers and types of organizations, degree of integration, structural forms, vertical and horizontal linkages, stability, coupling, crises, vulnerability to attack). Meyer described mental health organizations as highly "decoupled" from formal legitimacy structures due to complex and rapidly changing environments:

> Formal structure becomes in part a masquerade, and detailed control linkages between it and activity are avoided. ... Nothing is ever quite clear. There are glowing or critical accounts of programs and policies, but what actually happens is left very muddy. Therapies, management programs, patient counts, and so on are described, but in ways that do not make clear who is doing what to whom. ... The obfuscation is necessary and reasonable, created by people trying to protect some order in a conflicting and inconsistent environment ... It takes very able and thoughtful people to write and formalize descriptions of a mental health organization and treatments that cannot be understood (1985: 598; emphasis added).

Meyer's comment suggests that we can not hope to substantially improve the mental health system without subjecting to scrutiny the implicit or untested theoretical narratives that lead to an assumption of system rationality.

Black (1986) presented an "Institutional Context and Strategy" framework for the study of mental health policy. While this article is clearly influenced by GST, Black substituted words like "sector" for "system" and "structure" for "subsystem". Black, a sociologist, emphasizes several impinging vertical and horizontal sectors in order to explain policy, in notable contrast to a traditional over-emphasis in political science upon formal institutions and on narrow linear top-down policy making processes confined largely to those institutions. The elements of his framework are dynamic, influenced but not determined by structure. "Strategy" provides the basic concept for the researcher to enter and understand the abstract system being created.

Black introduced his case study with a conventional scenario for the birth of deinstitutionalization which, like much of the literature, over-emphasizes the cost and fiscal factors which are conveniently visible and immediate. His conception of policy as the "dependent variable" is responsible for an overly linear causal explication. Black's framework posits elements as "dynamic and nonrecursive, in that causality is interactive and moves in all directions, much as in an open system" (1986: 239). Given such assumptions, it is inconsistent with GST epistemology to conceive an event as a "dependent" variable, explaining it by tracing a backward path in time. This means trying to understand a large complex system, dynamic in history, from the direction of a more concrete manifestation of that system, rather than pursuing the synthetic approach to analysis: dialectically filling in the relations between system and its manifestations. Not only does Black's backtracking explanation provide a distorted view of the system, it can misleadingly concretize an "event" as a product at a moment in time—when it is rather a process over time.

However, Black uses a more synthetic approach when he describes the New York State Community Support System. He had the benefit of prior work with colleagues in comparing the 50 American states with respect to many structural characteristics, policies, and outcomes. Black's application of his framework to the case study reveals many interesting interactions and strategies that make the outcomes understandable, even if they would have been hard to predict.

Targeting Structural Constraints. Aside from whether this or any research can improve our ability to make non-trivial predictions in the social sciences, Black's approach—which directs us to issues of strategy and dynamics in multiple contexts—enriches the field of mental health policy and politics. Black's conclusions can be summarized by his comment: "If states could develop policies that would enable them to become less dependent on dominant structural interests ... their innovative policies in community mental health would stand a greater chance of success" (1986: 272). This implies a fundamentally different strategy in planning mental health reforms:

> As opposed to technocratic planning which merely inserts new structures within existing social and economic contexts, a political strategy would recognize the power and interests of the various actors within the system with a view to introducing or enhancing constraints or incentives to facilitate policy implementation. At this stage it appears that the emphasis should not be on mental health policy in a narrow sense but on the policy environment (McCubbin 1994a: 47).

Insufficient size in the system studied (in terms of historical period, actors, groups, institutions, and dimensions) can lead the analyst to play down even the subsystems which the analyst wishes to understand, since system inputs are affected by subsystem outputs to the system's environment. For example, it is rarely considered how overt behavioural effects of psychiatric drugs and drug withdrawal have contributed to the public's perceptions of the manifestations of mental illness, and hence to pressure to do something about "cracks" in the system (see Cohen D. 1997).

Bernstein and Lennard (1973) are among very few authors who have used a systems approach to model the context of prescribed psychoactive drug usage (outside-in approach) and to also address the impacts on society of such a system (inside-out) [see also Cohen D. 1988; Cohen D. and McCubbin 1990]. Bernstein and Lennard (1973) addressed implications at the levels of therapy, iatrogenics, social behaviour, and creation of powerful vested interests. The article details interrelations among systems affecting drug use, showing physician gatekeepers in the centre of a traffic circle receiving inputs and providing outputs to an environment of interacting systems of public policy, social forces, and community and group, as well as the personal and interpersonal domains. The authors discuss some of the social dynamics at work maintaining and increasing the use of drugs as "solutions" for social as well as personal problems.

Bernstein and Lennard distinguish between viewing a practice at the micro versus the macro levels, and are sensitive to the importance of finding the appropriate level(s) for research and solutions:

> An ecological model can show how physiology, society and nature can be integrated into a large pattern of interacting systems in delicate balance, demonstrating clearly that interventions at any level disturb the balances and the relations among other parts of the system. One needs to examine the effects upon the whole system from each input. It is one thing to expect a physician to assess the physical effects of a given dose of a particular drug upon a single person; it is quite another matter to rely upon physicians to assess the social consequences of maintaining 500,000 persons on methadone, another million on antipsychotic drugs and untold numbers on a variety of other potent biochemical agents (1973: 17).

The authors' point has not been heeded; the public and the legal system still defer to medical expertise with respect to <u>social</u> health policy questions. Evaluation of system-wide causes and consequences of an aggregation of multiple micro-level <u>medical</u> interventions is not <u>ipso facto</u> a <u>medical</u> question. The mental health system as a whole is not only defined by its subsystems but also by its suprasystems and by interactions within and between levels.

<u>Strategic Planning.</u> Failure to acknowledge the interactions of the studied system with impinging parallel and suprasystems is bound, particularly over time, to create a highly distorted image of the studied system. Casey et al., reporting on the views of participants at a symposium on the interface of the justice and mental health systems, noted such an effect:

> Although the actors in each system may understand their particular roles and responsibilities, few, if any, have an understanding of the full system—or even what happens in the system immediately prior to or following the completion of their duties. The fragmentation of services and funding arrangements across systems only adds to the difficulty of trying to understand the "big picture" (1992: 115-116).

In response to those concerns the symposium aimed to create an agenda for reform "based on a systems paradigm rather than the traditional doctrinal analysis paradigm" (1992: 107). Focusing upon system interactions and interfaces necessarily highlights dynamic system change within a broader context: it is within that context that similarities and divergences between the systems (system objectives, professional ideologies, public expectations and attitudes, regulatory mechanisms) can be discerned. The symposium participants suggested that too often reform attempts were made in a vacuum, without regard to systemic contexts, and hence too often inefficacious. They also noted, however, that there was so little in the way of feedback to policy makers of the results of what were generally ad hoc incremental policy changes, that there was no basis to monitor their successes or failures and, importantly, to further adjust policies to progress toward objectives.

This "Nominal Group Technique" symposium viewed its agenda as a dynamic "work in progress" (Casey et al. 1992: 109; as such, structured rather like Shakun's [1981a, 1981b] "inquiring system" for policy development). The dynamism of this process lay in its search to create the information needed to develop better policy strategies, given important tendencies outside the system's formal institutions (e.g., demographic changes to age and race distributions of populations and clients), and subsequently adjust those strategies in response to feedback. The technique is also developmentally dynamic (Smith 1973): it not only aims to influence policy with better information on client makeup and needs and on how the mental health and justice systems actually operate mutually, but also develops proposals to change the power structures surrounding those systems (e.g., empowerment of system constituents through funding for lobby groups and other means).

While ambitious, this agenda was quite immediately practical: in recognition of the many systemic levels, its strategic plan involved a strategy for "small wins" (Casey et al. 1992: 126): pragmatic suggestions for interventions at the operations level, some of which would be relatively easily to implement. However, rather than generate a multitude of unrelated ideas, participants developed proposals within the context of explicitly stated objectives and, as such, the possible interlevel impacts of even minor, service level changes could be traced to higher levels. Hence, for example, research aimed at generating feedback such as comprehensive needs analysis, and formative and summative evaluations (particularly from client perspectives), could be expected to influence legislators and the public, as well as clients and professionals within the delivery systems.

Mental Health System Dynamics

The above discussion leads to the suggestion of an epistemological approach to the conduct of science implying subordination of the object of interest to a larger object of study. The researcher retains the object of interest almost in peripheral vision while focusing on the system(s) which give it meaning[3]). This approach needs to be distinguished from that which has become traditional in recent years, that of viewing the object as "contextualized in" or "structured by" the environment. Explanation of a system within "environment", as opposed to within a suprasystem, leads to difficulties in viewing the system of interest as dynamic (Buchanan 1972). While that system may receive and process inputs, the outputs are analytically meaningless in the study of the system, because there is no mechanism for showing how the suprasystem responds to the outputs of the system of interest with future inputs to that system. By definition, once one begins to show how system outputs are transformed and fed back, one has begun to model a suprasystem.

Explanation. In theory, in some instances a suprasystemic perspective may be able to provide a relatively simple explanation of complex subsystem phenomena, providing a kind of "fractal" (Gleick 1988). In the human sciences the concept of fractal is useful as a heuristic device for simplifying—albeit in losing detail—our grasp of complex human processes. For example, a series of incomprehensible or seemingly random or irrational actions at one level might become comprehensible when seen as tactics within a strategy formulated at a higher level (see Crozier and Friedberg 1977). It is usual for GST influenced researchers to seek or create hierarchies in order to simplify perceptions of a reality which is far too confusing

---

3. Or, incorporating a more teleological approach: the systems which, when the object is situated in them, manifest meaning. See: Bateson, "Form, Substance, and Difference" (1972: 448-466); von Bertalanffy, "The Relativity of Categories" (1968: 222-250); Foucault (1966).

to learn from. In the mental health policy sciences, however, we are in danger of looking so hard at the trees as to miss the forest. Rochefort's review of the literature concerning policymaking cycles in mental health revealed such an enormity of intervening variables in models of mental health system change that "analysts stand at risk of 'overdetermining' the phenomenon under study" (1988: 148). Rochefort noted that none of the cycle theorists are prepared to use their models of system change (or policy change) for prediction.

In our view this suggests that fundamental, high level variables behind system change have not been identified or operationalized, or that interaction processes among variables and subsystems (requiring enclosure within a system) have not been adequately modeled. It is too easy to make an argument that mental health system dynamics have been, under Rose's typology of policy change models (cited in Rochefort 1988), any of static, progressive, discontinuous, or cyclical, by the selective use of criteria, facts, and time periods.

Models of system dynamics need to be put to predictive tests in order to provide some indication of their relative explanatory abilities; otherwise they may at most describe what the researcher has chosen to perceive. Admittedly this is far easier said than done, particularly insofar as hypothesized system dynamics could take centuries to observe in the future. However, researchers can "test" models by predicting the past. The researcher needs at some point to display a Foucauldian "contempt for facts" by creating a model which is not only intended to fit the facts at hand. The researcher does so by selecting key variables (rather than events) and the theoretical processes interrelating them, which can be subsequently tested upon more facts than were at hand when the model was created. At least initially, such models need not fit all the facts, but they should be simple, theoretically well described, and susceptible to elaboration or adjustment in order to accommodate new theory and types of facts.

Whether a system is studied as an "open" system, whereby the environment is regarded as providing exogenous forces and perturbations, or rather understood in the context of suprasystems, will determine our appreciation of the object system's dynamic behaviour. In other words, while we might be able to explain an open system's homeostatic or adaptive changes—possibly modeled as periodic or cyclic given environmental influences which include policy initiatives (as done by Morrissey et al. 1985, with respect to the mental health system)—we are unlikely to be able to explain heterostatic change whereby a new system emerges, perhaps suddenly, out of an old system which is radically different in kind from its later manifestations. We can be easily misled if we try to explain in adaptive or cyclical terms the relatively sudden emergence of a system of mental health characterized largely by a vast number of individuals consuming psychotropic drugs "in the community", from an earlier system serving fewer persons (but still many) residing for long periods of time in large asylums (McCubbin, 1994a).

Indeed, we may not even be able to discern such dramatic change, since what is "important" about a system is a matter of observer stance (Thompson 1977): a child on a raft who sees nothing but ocean may feel she is homeostatically rising and falling, rather than heterostatically going somewhere. Perception of direction requires an enlargement of the frame of reference (knowledge of ocean currents, use of binoculars to view land-masses, or simply faith). Scott's article "Systems Within Systems"—virtually the only systematic application of GST to the mental health system as a whole—recognizes this, identifying the "societal sector as an important system level of analysis":

> Thus the sector of mental health would include not only the collection of diverse facilities that supplies mental health services but also the manufacturers of psychopharmacological agents, the educational institutions that train mental health personnel, the professional associations and unions, the mental health advocacy organizations, the financing and regulatory bodies, and so on. (1985: 609)

Scott referred to most systems theory as "relentlessly ahistorical" (1985: 605); he did not provide examples, but his subsequent reference to cybernetic systems suggest analyses inconsistent with modern systems theory approaches in the human sciences. Cybernetic analyses—which focus on closed systems—and narrowly defined structural-functional analyses can fail to incorporate the essential dimension of history. If we inquire from the perspective of an appropriate suprasystem we cannot help taking an historical perspective, since the dynamics of a subsystem—seen from its suprasystem—can only be comprehended by overt reference to the passage of time, whereas a characteristic homeostatic pattern of a cybernetic or fixed purpose system can be encapsulated in time—thereby excluding history from the analysis.

Scott suggested that the "developmental history" of the mental health system calls for reference to the capitalist economic system and "the developmental state of the larger systems in which they participate: community networks, state systems, associations, and the nation state" (1985: 606). His article briefly discussed the import of these suprasystems, providing a model for further elaboration and operationalization. Attention to the system's developmental history facilitates the identification of new trends or meanings—e.g., of "marginalization" (McCubbin and Cohen 1996; see Twaddle 1996).

Boundary. Scott admits that the boundaries of the mental health sector are ill defined and porous. It is a virtue of the systems approach that the writer and readers are drawn consciously to the issue of boundary definition. Any theory creates systems and boundaries. GST makes this process explicit, encouraging clear thinking and discussion of the boundaries which are universally "constructed" by the researcher rather than apprehended in reality. A critique of GST could be leveled by a positivist, who might consider the theoretical attribution of systemness as arbitrary, abstract, or invalid as not directly observable. Rome drew attention to the "unreality of conceptual boundaries" in the context of perceptual hierarchies: "The distinction of ostensible difference that typifies appearance at molar levels disappears with a shift to another level of the system" (1969: 324).

Awareness of boundary definition problems could help to avoid superficiality in characterizing the dynamics of the mental health "system", particularly over a long period of time. It is essential to identify whether we are concerned with legal provisions, housing methods, treatment modalities, or with what may be more fundamental variables such as tolerance of deviant behaviour, quality of life, liberty, empowerment, normalization or community integration. This must be done in order to begin answering the questions raised by historians, social control theorists, and cycle theorists, as to whether the "system" has made progress, whether the same ideas are continuously recirculated as swings of a pendulum, whether any "reform" inevitably becomes coopted by its opponents, and so on.

Dynamic Models of Dynamic Systems. An empirical model conceived within a systemic perspective can be built by progressively adding assumptions about the postulated system. Some of these assumptions have generic types, in that they have been discussed or defined for species of abstract systems by GST theorists (e.g., structures, processes, degree of openness, complexity). Assumption of system "types", in conjunction with other theory, speculation, and observation, implies expected system processes. Among the prime requisites for the efficacy of such a model would be not only its ability to model a heterostatically dynamic system, but to be so itself. Hence it would need to be robust, so that a change in one piece would not cause the whole edifice to fail (see, e.g., Ellencweig 1992).

The idea of self-organization leading to emergence is a crucial element in understanding system dynamics from a suprasystemic vantage point. In this optic, the mental health "system"—or those fundamental aspects of it which underlie the researcher's interest—cannot be expected to continuously follow over a long period of time any kind of smooth path, whether static, linear, exponential, spiral, etc., since increasing complexification implies emergence of the system into a new system different in kind. Such a system is heterostatically dynamic. The nature of the path ahead of it can then become totally different, bearing no relation to that behind. Understanding of this is can be enhanced by looking at how familiar concepts underlying narratives about the mental health system (e.g., "mental illness", "treatment", "patient", "community", "needs", "risk") have radically changed over time.

The mental health system and its suprasystems are always in a process of dialectical change, which of course also incorporates the researcher as an active agent. Hence researchers need to periodically rejuvenate their basic understandings of what the system "is", and redefine the objects of study and interest, as the system, suprasystems, and the researcher all evolve together in complex interactive processes.

Orientations for Strategic Reform

Ackoff (1974) suggested that it is better to start with a poor solution that improves than a good solution that gets worse; this perspective of a "solution" as a dynamic systemic process provides an alternative to the "rationalist" economics-influenced obsession with optimizing results for discrete decisions. To this end, a number of "evolutionary" processes have been advocated, e.g., Smith's (1973) incorporation of learning in participatory planning, Gilbert's (1995) integration of evolutionary theory with the biopsychosocial approach to human problems, and robust planning processes that can cope with suprasystemic (including cultural) change (Motloch and Woodfin 1993). Shakun (1981b) described an "evolutionary systems design" which shows how a policy making system can change not only itself but its environment. Fisette (1986) noted that "systémisme" in planning has a much more aggressive approach to the environment than does Taylorism—rather than react to "perturbations" from the environment the systems planner creates them to control their results and reduce uncertainty.

It has been rarely discussed how mental health system outputs alter the environment—with the result that subsequent system inputs, and therefore the system, become altered in kind. Thus, if a "reform" does not satisfy demands giving rise to it, we cannot assume that those demands will remain in the system, hidden or otherwise. For example, is it possible that the broad humanitarian ideals which contributed to deinstitutionalization and community care have weakened due to disillusion resulting from the failure of this attempt at large-scale social engineering (as with many others)? Policy interventions typically have social effects much wider than their supposed targets, altering distributions of power and stakeholders' perceptions of their interests (McCubbin 1994a).

Policy Levers and Triggers. Thompson's application of Vickers' "appreciative systems" to a model of U.K. National Health Service policy making showed how "regulation" of the system, to maintain stability in its environment, can result in changes to itself and the environment, or in influencing the problem definitions posed by policy makers in the environment, thereby hindering substantive change even though societal appreciative or value systems might have changed. Thompson described how the presentation of policy proposals, without "any pretense at objective evaluation", could trigger mental health system changes, when there had been underlying changes in societal value systems: "... such proposals not merely constitute a new fact in the situation, they create a new situation. Mere proposals irrevocably change the status quo, whatever the policy decision" (1977: 401).

It is important to be sensitive to the distinction between triggers and fundamental conditions. Fundamental conditions may facilitate change when new factors are added in, but may not be enough to sustain reform objectives in the long run as the nature or importance of the new factors alter—which alteration may be partly a function of the changes induced by them. Hence, while cost and fiscal factors can be seen as important triggers to the deinstitutionalization process, how that process has unraveled needs to be understood in terms of two fundamental dimensions: social/cultural changes reconceptualizing "mental illness" and the place in the community of persons experiencing emotional distress, and the political/economic structure of the society bound to mediate any attempt to implement reforms pursuant to the social/cultural dimension.

It has become increasingly clear that the actual results of reform initiatives triggered by cost and fiscal factors have reflected more than a need to reduce spending in the mental health field, since, because of concerns that institutionalized persons be treated more humanely, asylum spending may not have greatly decreased. Indeed, spending on medical services and psychoactive drugs has greatly <u>increased</u>, reflecting the influence of the psychiatric profession and the medical establishment, as well of pharmaceutical companies and family lobby groups. This process was described within a systems perspective by McCubbin (1994a), who proposed that disillusionment and policy paralysis be replaced with a teleological approach to planning: a long-term strategic plan based on humanistic goals and continuously corrected, which emphasizes the policy <u>environment</u>. That article illustrated how even well-intentioned "rational" policy, ostensibly favouring clients, can be distorted in a program field structured by social/economic subsystems characterized by complex distributions of power, incentives, and constraints.

<u>Values and Planning.</u> The problem of modeling the mental health system with the aim of finding ways out of such reform impasses was pursued by McCubbin and Cohen (1996), who provided an implicitly systemic political economy framework for assessing the rationality of the mental health system. The proposition that psychiatrists can be expected to act in the interests of patients was challenged on a number of theoretical grounds, and patients were described as facing severe barriers in access to power within the system. It was concluded that psychiatrists cannot be expected to represent the interests of clients on the basis of the service, best interests, or medical models, and that client groups have systematic

disadvantages due to the therapeutic context, the stigma of mental illness, disabilities resulting from the illnesses and their treatments, the "free rider" problem, and drug company funding of psychiatric associations.

In such circumstances we can understand how progressive reforms based on widely-shared values—even if shared by actors also having interests conflicting with reform manifestations—can be unsuccessful. Individual actors share not only social values but also those related to their individual circumstances which can partly be understood in terms of their situations within systems of professional and economic organization. Furthermore, the actions and interests of <u>groups</u> are not always reducible to the circumstances and values of individuals making them up—as both sociologists and systemic thinkers have long emphasized.

In recognition of this, McCubbin and Cohen (1996) in effect stated that major progressive reform in the mental health system cannot rely upon the good intentions of paternalistic governments and caregivers. The incentives and constraints which structure systems and the behaviour of actors must be expected to enter into the equation as well. Therefore, given such a playing field, it becomes crucial to explicitly address the distribution of power in the system. Once we address the question of who has power, it becomes very obvious that even though the system is supposed to be centered around user "needs", users themselves have so little power compared to other actors as to amount to none at all. Therefore, that article suggested orienting policy attention towards clients as policy <u>agents</u> rather than as policy <u>objects</u>. Such a transfer of power is justified not only on "empowerment" grounds of therapeutic benefit (Hagan and Smail 1997; Winick 1994), but also due to the very pragmatic reason that the objectives of a reform will best be met in the long run by placing control of the shaping of the reform in the hands of those whose interests converge most with the reform objectives.

<u>Actors at Multiple Levels.</u> The above discussion illustrates the fundamental principle of complex systems: we cannot understand the whole without understanding the parts, nor the parts without understanding the whole. The analytical approach outlined here reflects that maxim by describing actors both as individuals having particular values and circumstances structured by systems at group, community and social levels and simultaneously viewing those systems

also as actors. Hence the dynamics of a system that incorporates both types of actors are very complex, and can not be predicted by simply considering each component as merely an "actor" with a singular obvious purpose (indeed, it should be realized that in human and non-human empirical situations "component" is a matter of definition, and no component can be shown to be indivisible into smaller components).

Even if each member of a group or category can be seen as having the same inherent values, the contexts of those individuals mediate those values, such that individual or group actions cannot be predicted by relying mostly on assumptions about individuals' values. For example, it does not necessarily follow that the actions of professional associations are designed to be in the "best interests" of clients, even if at the treatment level each professional may hold that guideline paramount. Furthermore, the task of predicting group interactions requires a dynamic, strategic analysis: strategic considerations can result in counterintuitive actions, and the context and patterns of the interactions can mean outcomes not anticipated by an essentially static analysis of the system (Crozier and Friedberg 1977; Fontaine 1997).

Strategic Planning within a Complex, Dynamic, and Power-Imbedded System

However, even once we recognize that implementation of user-centered community care requires user empowerment, how empowerment policy should be implemented is not obvious. While postmodernists, left-wingers and right-wingers have been converging upon the view that public policy should increase public participation (Boguslaw 1979; DeLeon 1994; Dobuzinskis 1992; Rappaport 1990), and while stakeholder participation in a well-designed process can improve the resulting decisions (Charlton 1991; Shakun 1981a), the dynamics of such decision systems, falling as they do within broader systems of power, can easily have unfortunate results. One may imagine a decision system designed to enhance particular interests, while increasing the participation of losers and thereby coopting them; or a system which enhances its legitimacy by broadening and strengthening stakeholder participation yet excluding ("justifiably", "unavoidably", or "inadvertently") certain classes of marginalized persons who have the most at stake but are the least recognized (see Charles and DeMaio 1993; Emanuel and Emanuel 1997; McLean 1995; Milewa 1997).

Hence planning for major long-lasting reform of the mental health system requires systematically incorporating the multidimensionality of the mental health system and suprasystem(s) into a framework which emphasizes interactive processes over time, rather than deterministic forces in sets of static snapshots (e.g., White's [1996] application of strategic incrementalism). If such a model could be created we would have a basis for beginning strategic planning to achieve meaningful reforms that have so far eluded us. Given the barriers to reform already highlighted by the literature discussed earlier in this article, viewing the mental health sector as a complex, dynamic system structured by political, economic and cultural dimensions, a strategic plan should start with attention to professional systems, to the issue of mental health administration as a subsystem of "health" (the more exact term in present circumstances would be "medicine"), to the pharmaceutical industry, to the interactions between psychiatry, drug companies, and universities, to the commonalities of mental health problems with other social problems[4], and especially to the ability of users to represent their own interests when they are often in conflictual competition with those of other actors.

Conclusion

There is a need for more integrative work—not simply to add more variables and dimensions to existing analyses, but to provide a coherent flexible framework for making sense of it all. Lack of a priori theory, and of an approach to method pushing the researcher "from behind", inevitably results not in an inquiring system

---

[4] Organized crime, drug dependency, anonymity and alienation, violence, abuse and neglect, environmental destruction, poverty, insecurity, cultural genocide and suicide, prisons, asylums, food banks .... Are all of these "separate" social problems, to be solved by experts in each domain? A systemic perspective might regard them as mutually reinforcing, and perhaps also as part of a larger malaise.

but rather an attracted system whereby the researcher is drawn to the available data and unconsciously adopts well-established perceptions. Breakthrough thinking need not be the result of "a complete contempt for historical facts" (Jones 1988: 86, referring to Foucault). However, it is unlikely to result from merely gathering and reporting the facts, unless that process is inspired by a vision that suggests what types of facts need to be generated or gathered, and how they should be organized and interpreted.

Implications of GST Epistemology

GST provides both a critique of the modernistic scientific paradigm and guidance to the researcher in approaching the structuring and interpreting of "reality". It is both philosophical and practical in its implications. It recognizes the enmeshment of the observer within the observed, making a virtue of this rather than leading to a stance of radical relativism. GST explicitly recognizes that research is, unavoidably, social action. Hence, the researcher should act responsibly in a dynamic social process in which research may help society to understand itself, its universe, its opportunities and threats, the meaning of fulfillment and how it might pursue it.

Certainly no systems theorist has yet captured the explanatory richness of a broadly defined mental health system over time as have some postmodernists, political economists, and social control theorists, like Foucault, Ehrenreich, Castel, and Szasz. These writers chose a very high level of perspective, in that they place understanding of the system in deep political, historical, economic and epistemological contexts. GST states that a system cannot be understood from a perspective entirely internal to that system: it is not clear that GST offers any advantages to understanding the broadest suprasystems enveloping the mental health system compared to other approaches concerned with the same level who have Marxist theory, "archaeology", or spiritual faith to draw upon. However, the process concepts of GST do offer the high-level thinker aids in modeling the dynamics of a subsystem, and draw attention to types of elements affecting and effecting change that might otherwise be missed.

Power and Epistemology. Overt application of GST to social problems has yet to reach its potential; its influence may have been hurt by the right-wing entrenchment seen in western countries since the 1970s. GST not only challenges modernism but encourages explicit attention to suprasystems, leading to recognition that power and economic privilege, rather than rationality, justice, and democratic ideals, often underpin the status quo (Twaddle 1996) and dominant discourses. GST is not, of course, the only challenger to modernism. Prominent schools of "constructionism" or "deconstructionism" share ground with GST in terms of its epistemological critique of modernist science. Some of them focus exclusively on social constructions of problems, narrative, and dialectic, as the objects of study and interest (Flaskas 1994). Although these studies note the results of power as expressed in symbols, in this conservative era they have tended to support the status quo by neglecting the power that includes or excludes persons and groups from the conversations that create those symbols (Forester 1995).

Such neglect may, in fact, express an implicit moral stance counter to emancipatory goals with respect to marginalized peoples. Prilleltensky, in an essay on the moral implications of psychological discourse and action, wrote:

> ... [M]ost postmodernists fail to offer an evaluative framework in which the merits of discourses can be assessed. They have decried universal frameworks that serve the interests of the powerful and neglect the needs and realities of specific people. This is an understandable concern, but it need not lead to the current reluctance to endorse any moral framework. ... [A] metacriterion, the need to explore and to express our moral sense, should remain firm, for being passive or recondite in the face of injustice is a moral choice that is hard to defend (1997: 528).

The most socially pervasive forms of power tend to become invisible, or unrecognizable as power. Widely accepted narratives—e.g. the discourse of science—can cloak power, providing a name for its exercise that distracts the eye away from the object of power toward the authority of the figure wielding it (Keen 1997). As Larner observed in his deconstruction of power in therapy, "power, like truth, is both a real construction and a constructed reality ... and on the scale of the universe, a reality that can vanish before our eyes" (1995: 213). An explicitly value based or theory driven stance can lead the researcher to systemic preconditions and constraints affecting what people experience and what choices they may make, rather than to what is being talked about in a media and literature which is itself subject to those same systemic constraints.

Ethics and Epistemology. Ironically, the image of GST has suffered from common misperceptions that equate it with extensions of modernism like linear programming, cybernetic studies of closed value-free systems, and management "systems analysis" aimed at implementing Taylorian ideals. General Systems Theory is not synonymous with applications incorporating assumptions of process linearity, equilibrium/homeostasis, closedness, deterministic causality, etc., although many disciplinary "systems" paradigms are, due to the influence of early cybernetic theory, unfortunately not refreshed with drinks at the well of interdisciplinary GST as it has continued to quietly develop into the 1990s.

An impression of GST as mechanistic is the opposite of the truth—rather, GST concepts ease the understanding and expression of ideas about humanity and the universe that underlie certain types of ethical stances. The "interconnectedness" of our universe suggests enhancing our ideas about social and ecological responsibility: the Kantian "enlightened self-interest". Our understanding of the capacity of complex systems to organize themselves and their environment suggests teleology rather than causal determinism, reinforcing not only our free will but our belief in it: liberating us but also driving home our responsibility to ourselves and others. Methodologies for simplifying complexity and developing strategy in a dynamic uncertain world offer at least the hope that through rational action—i.e., based upon strategies consciously calculated to advance and nourish our values—we might be able to progress toward the "good" that increased awareness and responsibility suggest.

Have we adequately explored how our treatment of those labeled mentally ill is a reflection of the power and interests of various social systems rather than a result of our collective ethical stance and of our potential ability to comprehend a richer meaning of mental health and illness and determine appropriate care? Do we, as individuals and as collectivities, perceive the "mentally ill" as them or as part of us (see Mossman 1997)? Do we understand the future implications of activities of our mental health institutions and practices, as they feed back to the mental health system through historical and social suprasystems? With respect to self-organizing systems and teleology, do the larger systems have adequate control of their subsystems? Is there danger of manipulation and cooptation from within? How can we use dynamic developmental ideas, explaining the change of complex systems, to create a program of care that improves due to its own internal dynamics—propelled by factors including the augmented agency of clients (see Swift and Levin 1987)?

Teleological Humanism

Explicit use of systems thinking leads to the recognition that any system is part of a larger system. Followed to its logical conclusion we are finally faced with fundamental ethical and metaphysical assumptions (Bowler 1981). Hence, analysis of the mental health system as a systemic inquiry must be critical, asking whether failures are essentially internal technical malfunctions of a system that generally functions well, or rather basic outcomes of a system which might, depending on the value-laden criteria applied, be considered to be poorly constituted (McCubbin and Cohen 1996; see Prilleltensky and Gonick 1996).

Innumerable efforts at reform of social policy systems has revealed a seemingly intractable paradox: "reform" <u>within</u> the system may be easy to instigate but unlikely to succeed; reform of the <u>system</u> might succeed, if only it could be implemented. For the GST philosopher the paradox is somewhat mitigated by the recognition that history is full of revolutions and sudden changes that would have been difficult to predict ahead of time on the basis of linear analysis of trends. While the growing recognition by mental health policy makers of the importance of listening to users has yet to translate into meaningful user empowerment, a variety of well-intentioned efforts have moved users a bit closer to the sites of actual power (Bassman 1997). And, as Dobuzinskis noted,

> ... when groups that have suffered from various forms of exclusion for a long time finally succeed in moving closer to power centers, sudden reversals of long established policies or new departures can be observed (1992: 363).

The most persistent of the social control critics have described how narrow-minded sociocultural or political economy forces are liable to divert major reforms based on ethical values. Nevertheless, there is reason to believe that our culture is rather tentatively and haltingly developing a humanitarian paradigm that is beginning to conflict with those regressive forces. A systemic perspective can help navigate through increasingly complex organizational and institutional systems that tend to reduce human agency or lend themselves to asymmetric (undemocratic, unfair) manipulation by those with societal power. Von Bertalanffy (1968) described possible consequences of increasing complexification as consistent not only with increasing intelligence and power but concurrently with the mechanization of subsystems and weakened control over them.

Strijbos (1994) called for the development of "systems ethics" for the health care system, noting that the dyad of patient-physician is now placed in organizational-technical collectives in which many other actors have responsibilities. Concern for systems ethics implies the belief that

> ... it remains possible, standing in our modern technological world, to break through to normative structures, which, however gravely violated they may be, remain present behind, and in, the actual structures, where they continue to hold sway (1994: 74).

A research agenda based on such an optimistic belief is justified by the increasingly urgent need to impose fundamental—hopefully humanistic—social values on our often monstrous social creations.

References

Ackoff, R. L. 1974. The Systems Revolution. Long Range Planning 7:2-20.

Anon. 1994. Inpatient, Outpatient Care Trade Dominant Roles in Mental Health Services Between 1955 and 1990. Hospital and Community Psychiatry 45:1052.

Appelbaum, P. S. 1997. Almost a Revolution: An International Perspective on the Law of Involuntary Commitment. Journal of the American Academy of Psychiatry and Law 25:135-147.

Austin, C. D. 1993. Case Management: A Systems Perspective. Families in Society: The Journal of Contemporary Human Services 74:451-459.

Bahg, C.-G. 1990. Major Systems Theories throughout the World. Behavioral Science 35:79-107.

Bassman, R. 1997. The Mental Health System: Experiences from Both Sides of the Locked Doors. Professional Psychology: Research and Practice 28:238-242.

Bateson, G. 1972. Steps to an Ecology of Mind. New York: Ballantine.

Becker, F. W. 1993. The Politics of Closing State Mental Hospitals: A Case of Increasing Policy Gridlock. Community Mental Health Journal 29:103-114.

Benson, P. R. 1994. Deinstitutionalization and Family Caretaking of the Seriously Mentally Ill: The Policy Context. International Journal of Law and Psychiatry 17:119-138.

Bernstein, A., and H. L. Lennard. 1973. The American Way of Drugging: Drugs, Doctors and Junkies. Society May/June, pp. 14-25.

Bertalanffy, L. von. 1968. General System Theory: Foundations, Development, Applications. New York: George Braziller.

————. 1981. A Systems View of Man. Boulder: Westview Press.

Black, B. L. 1986. Institutional Context and Strategy: A Framework for the Analysis of Mental Health Policy. In The Organization of Mental Health Services: Societal and Community Systems, ed. W.R. Scott and B.L. Black. Beverly Hills: Sage.

Bloche, M. G., and F. Cournos. 1990. Mental Health Policy for the 1990s: Tinkering in the Interstices. Journal of Health Politics, Policy and Law 15:387-411.

Boguslaw, R. 1979. Crisis in Social Science? Postpositivism and the Dilemmas of Social Planning. Society 16/5:40-44.

Boudreau, F., and P. Lambert. 1993a. Compulsory Community Treatment? I. Ontario Stakeholders' Responses to "Helping Those Who Won't Help Themselves". Canadian Journal of Community Mental Health 12:57-78.

————. 1993b. Compulsory Community Treatment? II. The Collision of Views and Complexities Involved: Is it "the Best Possible Alternative"? Canadian Journal of Community Mental Health 12:79-96.

Bowler, T. D. 1981. General Systems Thinking: Its Scope and Applicability. New York: North Holland.

Breggin, P. R. 1997. Brain-Disabling Treatments in Psychiatry: Drugs, Electroshock and the Role of the FDA. New York: Springer.

Brown, P. 1984. The Right to Refuse Treatment and the Movement for Mental Health Reform. Journal of Health Politics, Policy and Law 9:291-313.

Buchanan, J. M. 1972. Toward Analysis of Closed Behavioral Systems. In Theory of Public Choice: Political Applications of Economics, ed. J. M. Buchanan and R. D. Tollison. Ann Arbor: University of Michigan Press.

Buckley, W. 1967. Sociology and Modern Systems Theory. Eaglewood Cliffs, New Jersey: Prentice-Hall.

Campbell, J. 1982. Grammatical Man: Information, Entropy, Language, and Life. New York: Simon & Schuster.

Capra, F. 1990. Life as a Mental Process. Quest: A Quarterly Journal of Philosophy, Science, Religion & the Arts, Summer, pp. 7-10.

Casey, P., I. Keilitz, and T. L. Hafemeister. 1992. Toward an Agenda for Reform of Justice and Mental Health Systems Interactions. Law and Human Behavior 16:107-128.

Charles, C., and S. DeMaio. 1993. Lay Participation in Health Care Decision Making: A Conceptual Framework. Journal of Health Politics, Policy and Law 18:881-904.

Charlton, M. 1991. Controversial Decision Making—The Public Inquiry and Systems Thinking. In Systems Science in the 21st Century: Integrating the New Sciences of Complexity in Service of Humans and their Environment. Volume II: Advances in General Systems Theory and Philosophy, ed. S. C. Holmberg and K. Samuelson. Östersund, Sweden: Östersund University.

Cohen, C. I. 1997. The Political and Moral Economy of Mental Health. Psychiatric Services 48:768-774.

Cohen, D. 1988. Social Work and Psychotropic Drug Treatments. Social Service Review 62:576-599.

————. 1997. A Critique of the Use of Neuroleptic Drugs in Psychiatry. In From Placebo to Panacea: Putting Psychiatric Drugs to the Test, ed. S. Fisher and R. P. Greenberg. New York: Wiley.

Cohen, D., and M. McCubbin. 1990. The Political Economy of Tardive Dyskinesia: Asymmetries in Power and Responsibility. Journal of Mind and Behavior 11:465-488.

Crozier, M., and E. Friedberg. 1977. L'acteur et le système. Paris: Seuil.

Davidson, L., M. A. Hoge, L. Godleski, J. Rakfeldt, and E. E. H. Griffith. 1996. Hospital or Community Living? Examining Consumer Perspectives on Deinstitutionalization. Psychiatric Rehabilitation Journal 19(3):49-58.

DeLeon, P. 1994. Democracy and the Policy Sciences: Aspirations and Operations. Policy Studies Journal 22:200-212.

Diller, L. H. 1996. The Run on Ritalin: Attention Deficit Disorder and Stimulant Treatment in the 1990s. Hastings Center Report 26(2):12-18.

Dobuzinskis, L. 1992. Modernist and Postmodernist Metaphors of the Policy Process: Control and Stability vs. Chaos and Reflexive Understanding. Policy Sciences 25:355-380.

Dorvil, H. 1997. Reform and Reshaping Mental Health Services in the Montreal Area. International Journal of Social Psychiatry 43:164-174.

Durham, M. L., and J. Q. La Fond. 1996. Assessing Psychiatric Care Settings: Hospitalization Versus Outpatient Care. International Journal of Technology Assessment in Health Care 12:618-633.

Ehrenreich, B., and J. Ehrenreich. 1971. The American Health Empire: Power, Profits, and Politics. New York: Vintage.

Ellencweig, Y. 1992. Analysing Health Systems: A Modular Approach. Oxford: Oxford Medical Publications.

Emanuel, E. J., and L. L. Emanuel. 1997. Preserving Community in Health Care. Journal of Health Politics, Policy and Law 22:147-184.

Evans, R. G., and G. L. Stoddart. 1990. Producing Health, Consuming Health Care. Social Science & Medicine 31:1347-1363.

Fineman, N. 1991. The Social Construction of Noncompliance: A Study of Health Care and Social Service Providers in Everyday Practice. Sociology of Health & Illness 13:354-374.

Fisette, J. 1986. Les fondements pragmatiques de la planification. Revue internationale d'action communautaire 15:55-64.

Fitzpatrick, R., and D. White. 1997. Public Participation in the Evaluation of Health Care. Health and Social Care in the Community 5:3-8.

Flaskas, C. 1994. Postmodernism, Constructionism and the Idea of Reality: A Contribution to the 'ism' Discussions. Australia and New Zealand Journal of Family Therapy 16:143-146.

Foley, H. A. 1975. Community Mental Health Legislation. Toronto: Lexington Books.

Foley, H. A., and S. S. Sharfstein. 1983. Madness and Government: Who Cares for the Mentally Ill? Washington, D. C.: American Psychiatric Press.

Fontaine D. 1997. Repères pour une modélisation systémique des programmes de santé: Approach for a Systemic Modelling of Health Programmes. Social Science & Medicine 44:503-517.

Forester, J. 1995. Toward a Critical Sociology of Policy Analysis. Policy Sciences 28:385-396.

Foucault, M. 1966. Les mots et les choses. Paris: Gallimard.

Gilbert, P. 1995. Biopsychosocial Approaches and Evolutionary Theory as Aids to Integration in Clinical Psychology and Psychotherapy. Clinical Psychology and Psychotherapy 2:135-156.

Gleick, J. 1988. Chaos: Making a New Science. New York: Penguin.

Goldie, N., and L. Fredén. 1991. A 'Crisis' of Closure and Openness: The Present State of the Swedish Mental Health System in the Light of a Policy of Sectorisation. Social Science & Medicine 32:499-506.

Goodwin, S. 1989. Community Care for the Mentally Ill in England and Wales: Myths, Assumptions and Reality. Journal of Social Policy 18:27-52.

Haaijer-Ruskamp, F. M., and M. N. G. Dukes. 1991. Drugs and Money: The Problem of Cost Containment. Groningen, Netherlands: WHO Collaborating Centre for Clinical Pharmacology and Drugs Policy Science, University of Groningen.

Hagan, T., and D. Smail. 1997. Power-Mapping—I. Background and Basic Methodology. Journal of Community & Applied Social Psychology 7:257-267.

Hegarty, J., R. J. Baldessarini, M. Tohen, C. Waternaux, and G. Oepen. One Hundred Years of Schizophrenia: A Meta-Analysis of the Outcome Literature. American Journal of Psychiatry 151:1409-1416.

Hill, D. 1983. The Politics of Schizophrenia. Lanham, Maryland: University Press.

Hollingsworth, E. J. 1992. Falling through the Cracks: Care of the Chronically Mentally Ill in the United States, Germany, and the United Kingdom. Journal of Health Politics, Policy and Law 17:899-928.

————. 1996. Mental Health Services in England: The 1990s. International Journal of Law and Psychiatry 19:309-325.

Holstein, J. A. 1993. Court Ordered Insanity: Interpretive Practice and Involuntary Commitment. New York: Walter de Gruyter.

Hopton, J. 1997. Towards a Critical Theory of Mental Health Nursing. Journal of Advanced Nursing 25:492-500.

Ingleby, D. 1985. Mental Health and Social Order. In Social Control and the State, ed. S. Cohen and A. Scull. Oxford: Basil Blackwell.

Isaac, R. J., and V. C. Armat. 1991. Madness in the Streets: How Psychiatry and the Law Abandoned the Mentally Ill. New York: Free Press.

Jennings, A. 1994. On Being Invisible in the Mental Health System. Journal of Mental Health Administration 21:374-387.

Jones, K. 1988. Experience in Mental Health: Community Care and Social Policy. Newbury Park, California: Sage.

Juba, D. W. 1997. A Systems Perspective on the Introduction of Narrative Practice in Human Services Organizations in the Era of Managed Care. Contemporary Family Therapy 19:177-193.

Keen, E. 1997. Psychiatric Power and Levels of Concealment. Journal of Constructivist Psychology 10:125-137.

Kemp, D. R. 1991. Mental Health Crisis in California. Journal of Mental Health Administration 18:154-166.

Kenig, S. 1992. Who Plays? Who Pays? Who Cares? A Case Study in Applied Sociology, Political Economy and the Community Mental Health Centers Movement. Amityville, New York: Baywood.

Kiesler, C. A. 1992. U.S. Mental Health Policy: Doomed to Fail. American Psychologist 47:1077-1082.

Kleinman, D. L., and L. J. Cohen. 1991. The Decontextualization of Mental Illness: The Portrayal of Work in Psychiatric Drug Advertisements. Social Science & Medicine 32:867-874.

Lane, D. C., and M. C. Jackson. 1995. Only Connect! An Annotated Bibliography Reflecting the Breadth and Diversity of Systems Thinking. Systems Research 12:217-228.

Larner, G. 1995. The Real as Illusion: Deconstructing Power in Family Therapy. Journal of Family Therapy 17:191-217.

Lauzon, L.-P. 1993. Industrie pharmaceutique produits brevetés et prescrits. Une analyse critique: Rentabilité et situation financière. Exercises financiers 1990 à 1992. Unpublished research report.

Lurigio, A. J., and D. A. Lewis. 1989. Worlds that Fail: A Longitudinal Study of Urban Mental Patients. Journal of Social Issues 45:79-90.

Mack, J. E. 1994. Power, powerlessness, and empowerment in psychotherapy. Psychiatry 57:178-198.

Macy, M. W., and A. Flache. 1995. Beyond Rationality in Models of Choice. Annual Review of Sociology 21:73-91.

Mannheim, K. 1936. Ideology and Utopia. New York: Harcourt, Brace & World.

Marmor, T. R., and K. C. Gill. 1989. The Political and Economic Context of Mental Health Care in the United States. Journal of Health Politics, Policy and Law 14:459-475.

McCubbin, M. 1994a. Deinstitutionalization: The Illusion of Disillusion. Journal of Mind and Behavior 15:35-53.

————. 1994b. A Systems Theory Perspective of Mental Health Services and Policy Formation. Unpublished monograph.

McCubbin, M., and D. Cohen. 1996. Extremely Unbalanced: Interest Divergence and Power Disparities between Clients and Psychiatry. International Journal of Law and Psychiatry 19:1-25.

McLean, A. 1990. Contradictions in the Social Production of Clinical Knowledge: The Case of Schizophrenia. Social Science & Medicine 30:969-985.

————. 1995. Empowerment and the Psychiatric Consumer/Ex-Patient Movement in the United States: Contradictions, Crisis and Change. Social Science & Medicine 40:1053-1071.

Mechanic, D. 1996. Technologies for the Delivery of Mental Health Care. International Journal of Technology Assessment in Health Care 12:673-687.

Mechanic, D., and D. A. Rochefort. 1990. Deinstitutionalization: An Appraisal of Reform. Annual Review of Sociology 16:301-327.

Merrow, J. 1995. Reading, writing, and Ritalin. The New York Times, 21 October, p. 21.

Meyer, J. W. 1985. Institutional and Organizational Rationalization in the Mental Health System. American Behavioral Scientist 28:587-600.

Miedema, B. 1994. Control or Treatment? Experiences of People who have been Psychiatrically Hospitalized in New Brunswick. Canadian Journal of Community Mental Health 13:111-122.

Milewa, T. 1997. Community Participation and Health Care Priorities: Reflections on Policy, Theatre and Reality in Britain. Health Promotion International 12:161-168.

Morin, E. 1990. Introduction à la pensée complexe. Paris: ESF Éditeur.

Morrissey, J. P., H. H. Goldman, and L. V. Klerman. 1985. Cycles of Institutional Reform. In Mental Health Care and Social Policy, ed. P. Brown. Boston: Routledge & Kegan Paul.

Mossman, D. 1997. Deinstitutionalization, Homelessness, and the Myth of Psychiatric Abandonment: A Structural Anthropology Perspective. Social Science & Medicine 44:71-83.

Motloch, J. L., and T. Woodfin, T. 1993. General Systems Theory, Cultural Change, and a Human Science Foundation for Planning and Design. Systems Research 10:3-25.

Peirson, L., and R. Walsh-Bowers. 1993. The First Decade of the Canadian Journal of Community Mental Health: A Critical Review. Canadian Journal of Community Mental Health 12:23-35.

Prilleltensky, I. 1997. Values, Assumptions, and Practices: Assessing the Moral Implications of Psychological Discourse and Action. American Psychologist 52:517-535.

Prilleltensky, I., and L. Gonick. 1996. Polities Change, Oppression Remains: On the Psychology and Politics of Oppression. Political Psychology 17:127-148.

Prior, L. 1991. Community versus Hospital Care: The Crisis in Psychiatric Provision. Social Science & Medicine 32:483-489.

Rappaport, J. 1990. Research Methods and the Empowerment Social Agenda. In Researching Community Psychology, ed. P. Tolan, C. Keys, and F. Chertok. Washington, DC: American Psychological Association.

Reed, S. C., and D. A. Lewis. 1990. The Negotiation of Voluntary Admission in Chicago's State Mental Hospitals. Journal of Psychiatry & Law 18(1):137-163.

Rochefort, D. A. 1984. Origins of the "Third Psychiatric Revolution": The Community Mental Health Centers Act of 1963. Journal of Health Politics, Policy and Law 9:1-30.

————. 1988. Policymaking Cycles in Mental Health: Critical Examination of a Conceptual Model. Journal of Health Politics, Policy and Law 13:129-152.

————. 1993. From Poorhouses to Homelessness: Policy Analysis and Mental Health Care. Westport, Conn.: Auburn.

Rochefort, D. A., and R. W. Cobb. 1994. Problem Definition: An Emerging Perspective. In The Politics of Problem Definition: Shaping the Policy Agenda, ed. D. A. Rochefort and R. W. Cobb. Lawrence, Kansas: University Press of Kansas.

Rome, H. P. 1969. Psychiatry and Social Change. In General Systems Theory and Psychiatry, ed. W. Gray, F. J. Duhl, and N. D. Rizzo. Boston: Little, Brown and Company.

Scott, W. R. 1985. Systems within Systems: The Mental Health Sector. American Behavioral Scientist 28:601-618.

————. 1987. The Adolescence of Institutional Theory. Administrative Science Quarterly 32:493-511.

Scott, W. R., and B. L. Black, ed. 1986. The Organization of Mental Health Services: Societal and Community Systems. Beverly Hills: Sage

Shadish Jr, W. R., A. J. Lurigio, and D. A. Lewis. 1989. After Deinstitutionalization: The Present and Future of Mental Health Long-Term Care Policy. Journal of Social Issues 45: 1-15.

Shakun, M. F. 1981a. Formalizing Conflict Resolution in Policy Making. International Journal of General Systems 7:207-215.

————. 1981b. Policy Making and Meaning as Design of Purposeful Systems. International Journal of General Systems 7:235-251.

Skyttner, L. 1996. General Systems Theory: Origin and Hallmarks. Kybernetes 25: 16-22.

Smith, R. W. 1973. A Theoretical Basis for Participatory Planning. Policy Sciences 4:275-295.

Strijbos, S. 1994. The Individual and the Collective in Health Care: A Problem of Systems Ethics. Systems Research 11:67-75.

Sutherland, J. W. 1973. A General Systems Philosophy for the Social and Behavioral Sciences. New York: George Braziller.

Swift, C., and G. Levin. 1987. Empowerment: An Emerging Mental Health Technology. Journal of Primary Prevention 8:71-94.

Szasz, T. 1994. Cruel Compassion: Psychiatric Control of Society's Unwanted. Toronto: John Wiley & Sons.

Thompson, D. J. C. 1977. The Systems Model of Policy-Making and the National Health Service. Political Studies XXV:397-403.

Thompson, J. W. 1994. Trends in the Development of Psychiatric Services, 1844-1994. Hospital and Community Psychiatry 45:987-992.

Torrey, E. F. 1997. Out of the Shadows: Confronting America's Mental Illness Crisis. New York: Wiley.

Twaddle, A. C. 1996. Health System Reforms—Toward a Framework for International Comparisons. Social Science and Medicine 43:637-654(errata 43[10]:III-IV).

Tyhurst, J. S., F. C. R. Chalke, F. S. Lawson, B. H. McNeel, C. A. Roberts, G. C. Taylor, R. J. Weil, and J. D. Griffin. 1963. More for the Mind: A Study of Psychiatric Services in Canada. Toronto: Canadian Mental Health Association.

Upshur, C. C., P. R. Benson, E. Clemens, W. H. Fisher, H. S. Leff, and R. Schutt. 1997. Closing State Mental Hospitals in Massachusetts: Policy, Process, and Impact. International Journal of Law and Psychiatry 20:199-217.

Wald, P. M., and P. R. Friedman. 1978. The Politics of Mental Health Advocacy in the United States. International Journal of Law and Psychiatry 1:137-152.

Wells, J. S. G. 1997. Priorities, 'Street Level Bureaucracy' and the Community Mental Health Team. Health and Social Care in the Community 5:333-342.

White, D. 1996. A Balancing Act: Mental Health Policy-Making in Quebec. International Journal of Law and Psychiatry 19:289-307.

White, D., and C. Mercier. 1991. Reorienting Mental Health Systems: The Dynamics of Policy and Planning. International Journal of Mental Health 19(4):3-24.

Winick, B. J. 1994. The Right to Refuse Mental Health Treatment: A Therapeutic Jurisprudence Analysis. International Journal of Law and Psychiatry 17:99-117.

Wortis, J., and A. Stone. 1992. The Addiction to Drug Companies. Biological Psychiatry 32:847-849.

Ziegenfuss Jr, J. T. 1983. Patients' Rights and Organizational Models: Sociotechnical Systems Research on Mental Health Programs. Washington, DC: University Press of America.

Pergamon

International Journal of Law and Psychiatry, Vol. 19, No. 1, pp. 1-25, 1996

0160-2527/96 $15.00 + .00

SSDI 0160-2527(95)00028-3

Extremely Unbalanced:

Interest Divergence and Power Disparities Between Clients and Psychiatry

Michael McCubbin* and David Cohen*

It has often been noted in recent years that lasting reform of the mental health system has been very difficult or impossible to achieve. It is also obvious that the clients of the mental health system have always had very little power in influencing the direction of the system. We argue in this paper that these facts are related. It appears to us that a major unstated assumption reflected in mainstream mental health policy writings is that the mental health system is designed *for* the clients by expert practitioners and administrators, with input from loving families and a caring society. Instead, we believe that the system can be *expected* to impede reform based on client needs and will resist attempts to allow clients greater power in the designing of mental health policies and practices.

Our argument is based on two simple propositions. First, that the interests of clients diverge from the interests of other actors involved in the mental health system. Second, that the personal and political power of clients to advance their interests is small compared to the power wielded by other actors. For purposes of argument, we restrict attention to clients and psychiatrists as key actors within this system. A more complete political economy analysis of the mental health system would identify the interests, powers, and activities of all important actors within it including other health professions, families, public and private institutions, and drug companies.

---

*Correspondence and requests for reprints should be addressed to Michael McCubbin, Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec, Canada H3C 3J7.

A previous version of this paper was presented at the 20th International Congress on Law and Mental Health, Montréal, June 1994.

The authors wish to thank two anonymous reviewers of a previous draft for their helpful comments. Michael McCubbin would like to thank Prof. Harvey Simmons, York University, and Dr. Elizabeth Carlson, Ryerson Polytechnic University, for their advice and inspiration early in the preparation of this paper. David Cohen thanks the Conseil québécois de la recherche sociale and the Fonds de recherche en santé du Québec for financial support while this paper was being written.

Three major narratives model the psychiatrist/patient relationship as one in which the participants have convergent interests: the medical, market, and best interests models. We find that for various reasons none of these models justify a general assumption that the actions of psychiatrists will substantially meet client needs at the therapeutic and policy levels. Also, while it is evident that clients themselves have very little power and influence within or upon the mental health system, particularly compared to family lobby groups and professional associations, surprisingly few authors have systematically looked at the sources and extent of these power disparities, particularly as compared to psychiatrists. We attempt to fill this gap in the literature and hope that our analysis will facilitate the development of efficacious remedial policies.[1]

The Mental Health System's Resistance to Reform

If analysts of mental health policies agree on one thing, it is that the mental health system has been virtually intractable with respect to meaningful reform (Bloche & Cournos, 1990; Marmor & Gill, 1989; Mechanic & Rochefort, 1990). The system has changed its appearance in terms of institutional structures and laws, in terms of the jargon used in psychiatric diagnosis and etiological explanations, in terms of participation of nonmedical professionals, in terms of the more visible characteristics of treatment, and in terms of client populations. However, there is general agreement among both critics and supporters of the system that its *outcomes* have not improved, and in several respects have worsened (Breggin, 1991; Isaac & Armat, 1990). Commentators bemoan, on the one hand, the increases in the numbers of those considered mentally ill and the numbers of the mentally ill who are homeless or jobless, or on the other hand, the magnitude of treatment-induced illnesses and the extent of involuntary interventions.

Various factors have been identified in the literature as hindering reforms aimed at improving outcomes, but there is no consensus on which factors are essential. This is not surprising as there are widely diverging ideas about what mental illness is and how society should deal with those considered afflicted. For example, most psychiatrists today share a medical model orientation, part of a positivist view of science, that leads them to view mental disorders as diseases having primarily genetic or biochemical origins and requiring the same types of treatment as other somatic conditions. Psychiatrists' dissatisfaction with the current system stems from their belief that large numbers of mentally ill are not having their needs met because they are not receiving treatment for their illness (Mercier, 1989), because they or their representatives claim excessive legal protections (Torrey, 1988), because the state interferes unduly in therapeutic practice (Lamb, 1994), and because hospitals are underfunded

---

[1] In order to generalize, we have ignored several exceptions or extensions. For example, psychiatry is not a completely monolithic institution and many generalizations about that profession that follow do not apply to practitioners in full-time psychotherapy practice or to those who have adopted a holistic, preventive, or public health approach, that is, those whose diagnostic procedures and treatment plans incorporate as much as possible contextual factors involving family, general health and fitness, economic circumstances, community support, experiential background, spiritual values, etc. We have also ignored ways in which our arguments could be extended to all helping professions in the mental health field.

(Balter & Uhlenhuth, 1992). Some of these beliefs imply a clear distinction between what individuals need and what they say they want.

There is so little direct presentation of client views in the academic literature that the only generalization possible with respect to their views on reform blockage is that clients themselves have little power within the policy system (Chamberlin, 1990). Among the relatively few clients active in self-advocacy, two distinct perspectives have developed (Everett, 1994). The "consumers" accept the medical model but want increased choice and access to services (generally stressed in terms of private or public health insurance coverage). They hold that the major impediment to improvement is lack of public awareness about the prevalence and seriousness of mental illness and about the progress the medical and pharmacological sciences have been making and could continue to make given adequate or increased funding. On the other hand, the "survivors" mount various critiques, some of them in the form of fundamental challenges to the purpose and constitution of the mental health system. They typically share concern about forced treatment and other constraints on the liberties of individuals. They see reform as impeded by other actors in the system with more power than clients and with faulty ideas about clients' needs and the importance to the therapeutic process of patient sovereignty (Oaks, 1993).

In developing a narrative about the mental health system, authors are therefore hampered by the lack of agreement as to what constitutes the *needs* of the clients and therefore what a reform consists of. The literature can be seen as part of the system, reflecting the interests and ideologies of the writers (MacIntyre, 1973). It is spread out over professional journals from various disciplines, which may be why, with rare exceptions (see below), the *overall system framework* has not been directly addressed. The result has been a focus on policy initiatives rather than on the context within which policy is developed, and on countless explanatory variables that do not fit together coherently in a way that would allow for explanation rather than description. This leaves implicit and unexamined the assumptions, raised as problematic in the historical treatments by D. J. Rothman (1980) and Simmons (1990) that (a) major reform can be achieved, (b) it can best be achieved incrementally within the system, and (c) on the whole the system operates for the mentally ill. In another context, Wolin (1969) noted that the assumptions underlying a system frequently remain unexamined:

> If society is conceived to be a system of decision-making, and if the recurrence of unjust decisions is commonly acknowledged, it follows that the system is, to some persistent degree, a structure of systematic injustice, otherwise the idea of a system is an inadequate account (p. 1065).

The mental health sector is today almost always referred to as the mental health system. However, explicit use of systems thinking must lead to the recognition that any system is part of a larger system. Following this to its logical conclusion we are finally faced with fundamental ethical and metaphysical assumptions (Bowler, 1981, pp. 219–222). Hence, analysis of the mental health system as a systemic inquiry must be critical, asking whether failures are essentially internal technical malfunctions of a system that generally functions well, or rather basic outcomes of a system that might, depending on the value-laden criteria applied, be considered to be poorly constituted. So far, most systemic perspectives of the mental health system fall within the social control and "antipsychiatry" schools (e.g., Brown, 1985; Castel, Castel, & Lovell, 1982; Foucault, 1962, 1972; Fox, 1978; Scull, 1977; Szasz, 1991) and have been relegated to the fringe (Dain, 1989). Nevertheless, many of their observations, e.g., regarding conflicts of interest between those in the system with power and those without, remain the only fundamental analyses of the bases of the mental health system. Rather than evaluate the relative merits of competing paradigms about mental illness and its management, we choose to evaluate the policy playing field: even if clients, psychiatrists, academics, and other system participants disagree as to what clients need or how health care should be structured, are there not some *process* standards of fairness or efficiency which most actors could agree should apply with respect to policymaking?

Client-Centered Rationality of the Mental Health System

As discussed above, few commentators are satisfied with the mental health system; it has been variously described as being permanently in crisis, irrational, doomed to failure, incapable of meaningful reform. If true, such characterizations indicate fundamental systemic weaknesses: the system does not adequately achieve its purpose, insofar as that purpose is to relieve the distress of the mentally ill and help them to function within society.

We undertake to refute the implicit, widely accepted proposition that the system *will* likely meet client needs. We do so by demonstrating the proposition's internal logical weaknesses and by providing empirical evidence that suggests that its assumptions are not met or its predictions are not borne out. We begin by identifying those major arguments that might underlie confidence in the system with respect to its client-centered rationality.[2] Assessment of those arguments leads to our conclusion that the mental health system has no good theory to support the proposition that needs will be met, whether or not what clients want or say they want is what they need.

It is a simple matter to identify sufficient conditions for the expectation that the system will meet client needs if we assume a broad mental health system whose activities result from individual strategic decisions made in complex, uncertain contexts over time. These contexts are structured by socioeconomic patterns of interaction and exchange and constrained by distributions of resources such as money, influence, location—that is, *power* (Black, 1986; Crozier & Friedberg, 1977; Etzioni, 1982; Scott, 1985). We have identified three individually sufficient conditions supporting the hypothesis that the mental health system can be expected to meet client needs:

---

[2] This paper assesses the rationality of the system only with respect to the criterion that it meets the needs of clients. Hence, establishing that the system is irrational under that criterion does not exclude other grounds, such as protecting the public, maintaining cultural or economic norms, or providing jobs for health-care workers. To evaluate the system properly, it is necessary to separate conceptually different objectives of the system.

- *Interest convergence*. Powerful actors in the system have interests that converge with those of clients. Thus, sets of interactions in which actors pursue their own interests—however different from those of clients—are likely to serve the interests of clients.
- *Power access*. The system satisfies client needs as a result of client power to affect system outcomes—which power might be exercised in various ways throughout the system (at the political, policy, administrative, and therapeutic levels).
- *Policy intervention*. Insofar as the system does not meet client needs, policy interventions at the administrative and service delivery levels will gradually reform the system in that direction.

We will first review three models of the psychiatrist/client relationship that explicitly or implicitly underlie the proposition or assumption of interest convergence: the medical model, the market model, and the best interests model. We will demonstrate that the case for interest convergence cannot rest on these models due to their inherent logical weaknesses and constraints affecting their applicability. We then describe several factors that, in their combined impact, severely challenge the idea that clients are able to exercise power within the system in order to satisfy their interests or meet their needs.[3] These factors are: client role passivity, stigma, iatrogenesis, mental illness disability, difficulties in organizing self-advocacy groups due to the client "free rider" problem, and external funding asymmetries favouring nonclient groups.

With respect to the third statement in support of the hypothesis, the inefficacy of past policy interventions has been alluded to above. Several writers have described how past interventions have been co-opted as a result of the power of nonclient actors (McCubbin, 1994; White & Mercier, 1991). If it is true that actors holding interests incompatible with those of clients control outcomes *within* the system, it is also true that they have considerable influence over how the system is reproduced, that is, how the rules of the game are constructed and altered. Despite the impact of the growing client "consumer" movement, there seems to be few grounds for optimism (D. Cohen, 1994). Our final section will therefore clarify the areas to which policy interventions should be focused in order to improve the rationality of the mental health system. In general, we will suggest that policy should ensure that the underlying assumptions of the system's operative model(s) are explicitly recognized and satisfied. If not, the weaknesses of the model(s) should be either shored up by further interventions, or the model(s) rejected as a basis for policy.

Interest Divergence

The medical, market, and best interests models of the psychiatrist/client relationship are the major narratives variously used, implicitly or explicitly, to justify society's trust that the interests of psychiatrists and clients converge at the therapeutic and policy levels. We analyze each model to argue against this assumption.

---

[3] Our method does not prove that clients needs will not be met, but merely allows us to reject the hypothesis

The Medical Model

The term medical model (or biomedical model) has been used heavily for some 30 years, most often in critiques of mainstream medical practice, which assert that the implicit belief system underlying such practice is either too narrow (Engel, 1980), inappropriate to mental health problems (Szasz, 1991), or both. This model in effect views the human body as a machine that either functions properly or not. Medical practice under this model consists of identifying what is broken or worn in this complex machine and fixing or ameliorating it (Evans & Stoddart, 1990). However, the medical model, based as it is on a positivist paradigm (Frankford, 1994), has been seriously questioned in recent years. Systems theory (Buckley, 1967; von Bertalanffy, 1968), contextualism (Reese, 1991), chaos theory (Gleick, 1988), and other ecological approaches have shaken the atomistic, linear cause-effect paradigms underlying science until recently. The increasing use of alternative health practitioners (Buckman & Sabbagh, 1993), the demand of consumers for a wider variety of psychosocial and alternative forms of health care (Fisher, 1994), as well as the virtually complete rejection of the model by psychiatric survivors suggest that a mental health system based on the medical model may not be fully responsive to clients' interests.

In psychiatry, the failure to validate somatic pathology as the cause of the emotional distress encountered by psychiatrists further suggests that the medical model does not provide a sufficiently solid scientific justification for paternalistic interventions on clients considered mentally ill. For example, Szasz (1991) has written that mental illness is a myth in that it is a label applied to deviations from social norms rather than to observable abnormal somatic conditions. His argument is useful in that even if we were to believe that most serious psychological distress is somatically based, psychiatric diagnosis remains inexact, invalid, and fuzzy (Kirk & Kutchins, 1992; Mirowsky & Ross, 1989a; Wakefield, 1992). As such, it is liable to be influenced by contextual variables such as socioeconomic and cultural differences between doctor and client and race and gender role expectations (Caplan, 1992; Littlewood, 1992; Loring & Powell, 1988; Wade, 1993).

However, the case against interest convergence under the medical model does not rest solely, or even mostly, on the above considerations. Interest divergence appears more clearly when it comes to treatment. A treatment option may hold out the hope of relieving some symptoms (although, for most mental disorders, the actual efficacy of somatic-based treatments compared to alternatives remains doubtful; see Fisher & Greenberg, 1989), but may have disadvantages—including side effects, financial cost, implementation effort, and transition cost (due to altering habitual routines)—and might be administered in a manner or in an environment that is oppressive or distressing (Mosher & Burti, 1989). It is to be expected that psychiatrists will be more aware of treatment effects on observable behaviours that are relevant to diagnostic criteria than of effects on client experience, which the client observes more acutely than the doctor and may value quite differently (Brown & Funk, 1986).

---

that they will be, thus creating a reasonable *prima facie* argument for the null hypothesis (and main argument of this paper) that client needs cannot be expected to be met.

Furthermore, under the medical model the issue of client preferences need not arise (Butler, Campion, & Cox, 1992), as the only questions are scientific/technical matters totally within the purview of the medical practitioner. From the perspective of such a practitioner, the only question the client need answer is whether to be well or ill. Those who reject prescribed medical treatment are assumed to be choosing to be ill and are therefore considered irrational (Szasz, 1993).[4] They may be subjected by families or society to forced treatment. The medical model's claim to serve the public good lies in the implication that, except in what would be characterized as exceptional cases of practitioner abuse or incompetence, medical treatment affects only a single universally accepted criterion, i.e., reducing illness. Of course, like any abstraction, the medical model oversimplifies the basis on which mainstream medicine has been practiced. It has, however, been implicitly very powerful in guiding and justifying medical practice. Its association in the public mind with modern science has in the past helped to conceal medical decision making and protect it from interference by the public and clients (see Zola, 1978).

However, as "alternative" health practices achieve respectability and legal recognition, they often adopt much of the biomedical methodology (e.g., chiropractic: Coburn, 1993). Part of the reason for this is that medical interventions may be more visible and quantifiable, hence more suitable to the bureaucratic accountability and funding requirements of governments and insurance companies that too often have favoured tangible inputs rather than health outcomes (Morone, 1993). Furthermore, while the power of the medical model may have peaked in the treatment of somatic illnesses, in various respects it has been increasing within the mental health field. Kiesler (1992) notes that in the United States, mental health policy has been increasingly dominated by a health policy centered on hospitals, short-term acute care, and surgery. Public belief in somatic, especially genetic and biochemical "causes" of serious emotional distress and even of personality traits, nervousness, tension, and occasional depression seems to be on the rise. Within the psychiatric profession the influence of psychopharmacologists has risen, whereas practice of social psychiatry, dynamic psychotherapy, behaviourism and interactionism has declined (C. Cohen, 1993; Stein, 1990).

While more than a few mental health clients prefer a biomedical approach—especially given the respect and visibility it is accorded by society in general—it needs to be recognized that this approach is not explicitly directed to clients' needs or preferences, which for many go well beyond the symptom reduction (or replacement) typified by treatment regarded as successful.

The Market Model

The market model (also known as service, or consumer[ist] model) characterizes the encounters between health providers and clients as marketplace

---

[4] This is reflective of the "outcome approach," whereby personal decisions not reflecting community values are taken as evidence of incompetence. The appropriate values may be medical norms or the personal values of the physician: "Doctors like patients to agree with them. When a patient agrees with me, the patient is rational" (a psychiatrist in testimony at the Elizabeth Bouvia hearing, opposing her competency to choose death, quoted in Annas and Densberger, 1984, p. 571, footnote 39).

transactions between sellers and consumers (Butler et al., 1992). Services provided to customers in a competitive free market could be seen as efficiently serving the public interest, in accordance with the "invisible hand" theories of competitive markets (Smith, 1924).

In the rarely discussed last chapter of *Asylums*, about the "tinkering trades," Goffman (1961) provides a succinct and pertinent analysis of the applicability of the service model in psychiatry. He compares the psychiatrist/client relationship with his abstract view of the general server/client relationship and concludes that the former does not meet the basic assumptions of the latter. Few real-life service relationships meet the ideal, but the main point of Goffman's criticism remains that psychiatric services diverge far from a market ideal, wherein a consumer can shop for a service in a market where providers must compete to represent best the interests of consumers at low cost.

A large degree of market failure can be noted where psychiatric services are provided as market commodities. According to Goffman, factors militating against the idea that psychiatrists will act in accordance with their clients' interests include the differential selection and management of clients based on market-irrelevant variables such as social status and race; divided loyalties between client and the institution, where the provider works for an institution; the impact of professionalization whereby the provider responds to ideal rather than actual client interests; the triad of interests—client, helper, community—whereby "the more his own profession is given a public mandate to control [the client], the more [the doctor] is likely to be accorded the public task of maintaining community standards, which at times will not be in the immediate interests of a particular client" (1961, p. 339). However, the most important deviation from the service model in psychiatry occurs when the client is not fully consenting[5]: a distinction between "patient" and "client" thus arises whereby the patient is the object of medical intervention, but family or society becomes the client (pp. 344–345).

Goffman describes how various specialized professions in a large mental hospital are removed from direct responsibility to the client and act in accordance with ease and bureaucratic efficiency to the detriment of the service provider/client relationship. The relative importance of large mental hospitals has been declining, but this type of sociological criticism can also be applied to today's psychiatric wards in general hospitals and even to large community service organizations in urban centers (Prior, 1991; Regan, 1987).

Even if the market model were a realistic description of psychiatric care for the majority of clients, it cannot be invoked to justify treatment for those considered incapable of functioning as health consumers and certainly not for those subjected to mandated or *de facto* forced treatment. These two types of patient are neither clients nor consumers. The degree to which they are subjected to others' power and interests needs to be explicitly recognized before we can begin to find ways to protect their interests.

---

[5] Aside from the category of clients for whom treatment is mandated under criminal or mental health laws, impaired self-determination occurs in the far more numerous cases of client susceptibility to pressure or coercion due to the illness or society's perception of it (Diamond, 1995; Reed & Lewis, 1991; J. Rothman, 1989; Szasz, 1993).

In sum, in the mental health field, there is good reason to question the idea that market forces underlying the service provider/client relationship incite providers to act fully in accordance with the interests of clients. This is not to say, however, that the market model should not be used as an objective, subject to its limitations and other valid considerations, to enhance a helping profession's responsibility to the client and accountability to the public. There may be ways to alter the delivery of mental health services such that the market model applies to some clients. Basic criteria for market efficiency would have to be met: competitive rather than oligopolistic/monopolistic supply; informed consumers; and, where purchase of supply is subsidized, that demand is not excessive or distorted in its distribution among services. These conditions imply that incremental introduction of market structures cannot be justified without taking into account overall structures of the economy and of the health and social systems in place.

The Best Interests Model

The typical psychiatrist-client relationship cannot be characterized as providing straightforward technical expertise at the behest of the client to fix a specific, well-defined problem. Parsons (1951) described a professional model wherein the client is expected to trust the practitioner and the latter is expected to operate in the client's best interests. If the psychiatrist acts in the best interests of the client, it implies, by definition, using a degree of discretion or power in selecting a course of treatment according to the caregiver's judgment of what course would best serve the client's interests (Griffith, 1991). Under the best interests model, the *voluntary* therapeutic relationship would be idealized as the client authorizing the practitioner to fix or cure a "problem" and delegating to that practitioner decisions about the nature of intervention.[6] In this scenario, the client's lack of the knowledge and skills required to assess the probabilities of risks and benefits of treatment or of alternate courses of action prevent him or her from correctly making such decisions. In the *involuntary* context one would surmise that either the client is considered unable or unwilling to acknowledge or determine the nature of the problem or is unable or unwilling to delegate power to select the means to deal with it.

In the market model, outcomes are clearly chosen by the client, who is not expected to understand the means. The best interest model departs from the service ideal if the service provider possesses delegated or substituted decision making power regarding which *outcomes* are to be selected. A television owner who takes the set in to have the hissing removed does not expect to have it done at the expense of another outcome, such as colour quality. The difficulty in applying the market model to the therapeutic context lies in the complex relationships between means and outcomes. Few courses of treatment affect only one outcome, e.g., paranoia, without affecting others, e.g., alertness. The

---

[6] The fact that much, if not most, medical practice is still within the aegis of the medical model is revealed in discussions about whether to expand participation in decision making to clients. For example, England and Evans (1992, p. 1224) suggest "[d]elegating decisions to patients." But if clients own their own bodies, the issue of increased client participation in medical decision making should rather be seen as one of clients reducing their delegation of decision making to doctors.

market model, rather than the best interests model, better describes the situation where the mental health consumer selects a course of treatment after weighing the relative merits of potential outcomes of importance to the consumer, once they, and estimates of their probabilities, are explained by the practitioner. Therefore, to the degree that *informed* consent is provided by the client, the best interests model does not apply.

Appropriate use of the best interests model requires that all of the following three conditions be met:

- the provider is able to estimate with an expectation of reasonable accuracy the client's best interests;
- the incentive structure is such that this provider can be expected to try conscientiously to ascertain the client's best interests and exercise power or discretion in accordance with those interests; and
- there is justification for substituted or delegated decision making.

Ascertaining Best Interests. Use of the best interests model rather than the medical model implies that there are considerations in assessing treatment alternatives that go beyond the technical considerations of fixing something that is broken. The correct intervention aims to maximize net benefits to the client, taking into account the expected distribution of positive and negative treatment effects and how those would be valued by the client—a unique individual whose tastes are influenced by many situational variables such as culture, family status, and profession. This creates a very difficult decision problem for the delegated or substituted decision maker, further complicated by the need to consider the client's attitude to risk with respect to the probability distributions of outcomes (Gigliotti & Rubin, 1991). The estimation process itself is subject, therefore, to considerable error; the probable degree of such error reduces the likelihood that the expected net benefit of a substituted decision is positive (McCubbin, 1996).

Incentive to Act in Best Interests. For the best interests model to be an empirical reality rather than merely an assumption, we need to address whether psychiatrists possessing delegated/substituted decision-making power are observed to be (or can reasonably be expected to be) exercising their discretion in a conscientious attempt to determine explicitly the best interests of the client—a client noted above as having individual and situated values, tastes, and preferences. This will not be the case, for example, where the practitioner operates according to the medical model, a paradigm that does not call for a view of the client as a unique social individual.

Analysis of the incentive structure of psychiatrists yields a number of factors that should raise doubts about the reasonableness of the expectation that discretion will be exercised explicitly, rather than incidentally, in the best interests of clients. For example, professional negligence in a psychiatric malpractice suit can be defended against by merely establishing that the practice does not diverge widely from usual standards of care (recommendations put forth by professional associations, regulatory bodies, expert practitioners in the scientific literature, or simply that which is usually done by the practitioner's peers;

Simon & Sadoff, 1992). Court-ordered treatment generally relies on the "professional judgment" of the practitioner, despite attempts in the 1980s in the United States to establish a "best interests" or "substituted judgment" standard (Stefan, 1992). Over the last 30 years, increasing expectations and laws requiring informed consent may have widened client participation in decisions, but there has not been a parallel development of procedures to ensure adequate investigation and meeting of the client's interests insofar as decisions are made by the psychiatrist or other substituted decision makers (Griffith, 1991). In general, the norms of health institutions imply that treatment is indicated by the diagnosis rather than by the client's interests.

Ascertaining the interests of a client can be time consuming. None of the usual financial compensation mechanisms create a motive to ascertain best interests, although a psychiatrist paid by salary would be less disinclined than one remunerated by caseload, number of clients seen, or treatments provided. The psychiatrist is not paid or evaluated on the basis of health outcomes or client satisfaction. Rather, the health system requires simply that the psychiatrist be *seen* to be doing something. It would be difficult to monitor substituted decision processes involving relatively intangible acts, e.g., deciding *not* to treat, helping the client to formulate and express "the problem(s)," providing advice or assistance to obtain alternative or additional resources (housing, mutual support, spiritual guidance, social skills).

Most somatic treatments in psychiatry have well-documented adverse effects. Improvements from such treatments, especially where coerced, may be more apparent to nonclient actors, including families and society, than they are to clients, who may be more concerned with possible side effects than are doctors (see Donovan & Blake, 1992). In a detailed microeconomic analysis of the problem of selecting appropriate levels of substituted decision making, Appelbaum and Schwartz (1992) strongly suggest that review by clinicians is superior to judicial review; however, they acknowledge that they did not consider whether clinicians will defer to client preferences or instead impose their own (fn. 21, p. 449). Our discussion suggests that there are ample reasons to reject the general assumption that psychiatrists can be counted upon to ascertain the best interests of the client.

Justification for Substituted/Delegated Decision Making. It is commonly held that mental illness may impair the ability to recognize a problem and the volition to deal with it effectively. Psychiatrists are often trusted by their clients or by society to determine the nature of the problem. However, one of the main criticisms of the mental health system leveled by radical critics is that health-care providers have no right to substitute their judgment for that of the client where he or she is able to express his or her will. This position rests on two grounds: that one's ability to assess the need for mental health care is not necessarily impaired by psychological distress (Callahan, 1984), and even if it were, that it would be an assault on human dignity to provide treatment against the expressed will of an individual (Szasz, 1991).

One of the most interesting and controversial debates in the mental health policy field during the last 20 years has centered on the issue of involuntary treatment. Today, proponents of expanding legal powers of involuntary treatment advocate, in an Orwellian distortion, "the right to treatment" (Torrey, 1988). In its contemporary guise, this phrase means that treatment should be imposed on individuals considered incapable of acknowledging, due to mental illness, their own need for treatment. Although the generally accepted criteria for incompetence are normative notions about practical reasoning, their assessment is entrusted to psychiatric experts who frequently attribute incompetence to "lack of insight," operationalized tautologically as clients' frank disagreement with their treating psychiatrist regarding diagnosis or treatment (Cohen, Thomas, Dallaire, & Morin, 1995; Holstein, 1993). One problem with the right to treatment argument is the still-fragile nature of psychiatric knowledge: well-documented problems of diagnostic reliability and validity do not allow us to say with objective certainty that the client is irrational and the physician right when a decision is made to administer treatment against the will of the client (see previous section on *Medical Model*). Furthermore, the outcomes of treatments of the past and present, such as shock and drug therapy, are known to reduce some bothersome behaviours, but claims of cure have not been borne out, while systematic iatrogenesis is observed (Dewan & Koss, 1989). Hence, even where the client's medical incompetence is officially established, it does not follow that psychiatric treatment should be imposed against the *expressed* will of the client, or that responsibility for substituted decision making should rest with or be delegated to psychiatrists.

Many of the arguments advanced against involuntary treatment, particularly those that recognize a divergence between the interests and perspectives of psychiatrists and those of clients (Brown, 1984; Burstow, 1992; Cohen & Remler, 1992), can be found to support the proposition that the contributions of psychiatrists and their professional organizations to public mental health policy should not be taken to be necessarily in the best interests of clients.

Underlying Systemic Assumptions

Each of the models described above has some merit, given that theory can only hope to approximate reality, but in many situations their applications need to be more clearly circumscribed, reinforced, or replaced with alternate models. There is no assurance that a client's interests will be served where the psychiatrist operates under the medical model; a market form of health delivery, where it does actually exist, is inconsistent with meeting the needs of clients with reduced competence or subject to coerced treatment; and the difficulties inherent in conscientiously, accurately, and reliably assessing clients' best interests suggest that the outcome may fall far short of the expected. A mental health system based upon a haphazard blend of vaguely understood theoretical underpinnings is likely to fail: substantial reform of the system requires closer and much more explicit links between theory, values, and practice (see Moore, 1984).

Power Disability

That clients suffer some systematic power disadvantages with respect to psychiatrists at the therapeutic and policy levels is both obvious and rarely discussed explicitly. While some analysts have remarked briefly on the relative impotence of client groups (Mechanic & Rochefort, 1990; Simmons, 1990), most attention has focused on the efforts of individual clients and of client rights groups to achieve improved legal rights (Brown, 1985; Marmor & Gill, 1989). It is difficult to separate out the actual impact client rights groups have had in these efforts, given the convergence of interests in successful civil rights initiatives with those of other disenfranchised groups (e.g., implementation over the past decade of the Canadian *Charter of Rights and Freedoms* and of the *Americans with Disabilities Act*). In this section, several reasons are advanced to explain why the ability of clients and client groups to protect and advance their interests within the therapeutic, administrative, and policy contexts is systematically impaired with respect to the ability of psychiatrists.

Patient Role Passivity

One way to model the client's role in the therapeutic encounter is to identify its place along a passivity/proactivity continuum. The proactive client engages the services of the psychiatrist in order to fix problems defined by the client. This would be consistent with the market model version of the therapeutic encounter. Such a client would be assumed to be reasonably well informed about the therapeutic market and would monitor the practitioner's actions and the impacts of treatment to satisfy him/herself that these are in accord with his/her objectives. At the passivity end of the continuum, the patient is assumed to be unable to identify the nature of the problems and lacking in capacity to select a practitioner and choose a treatment plan. Treatments chosen by the practitioner (who is selected by institutions, family members, etc.) are applied *to* the patient. The success or failure of treatment is not for the patient to assess, but is the prerogative of others (Miedema, 1994). Hence, this type of therapeutic encounter infantilizes the client in assuming the client's lack of capacity,[7] and makes the client only the object of treatment rather than the subject. In this abstraction, the passive client's role is little different from that of the broken television set brought into the shop for repairs. The set has no choice as to the performance criteria to be set or the methods of repair. The patient, like the television, resembles a commodity rather than a client.

Numerous social structures and paradigms could result in producing the passive therapeutic context, e.g., slavery, feudalism, nonsecular fundamentalism, totalitarianism. Today in the Occident, however, the medical model provides the main paradigmatic support for the passive therapeutic context. Underlying the infantilization or objectification of the client, his or her problem (that is, his or her behaviour) is medicalized into a genetic, physiological or biochemical disease, beyond the control of the client; "such a belief leads inevitably to passivity, personal and political" (Hill, 1983, p. 435). These belief patterns are as likely to be held by clients as by practitioners and society (Kupers, 1993). Reduced power in the therapeutic context becomes isolation within the social/political context, since, as noted by Ingleby (1985, p. 163),

---

[7] "Medical practice has an 'infantilizing' effect . . . precisely because the powers it invokes are those of a parent" (Ingleby, 1985, p. 162).

"the medical model firmly locates the problem as residing *inside* the patients themselves . . . the illness itself can be treated separately from any environmental factors that may be seen as having led to it."

Since the (disordered) expressions of clients are often viewed by psychiatrists as precisely what needs to be modified by psychiatric intervention, they are devalued as inputs to the mental health policy process. The extent to which this is true is hard to measure directly, but the paucity of client perspectives in the mental health literature, and the small policy influence of clients and ex-clients compared with family members and psychiatrists begs an explanation. The clients themselves may feel devalued—feeling that their subjective experiences are not reliable or important. Furthermore, even those not holding such feelings may be discouraged from contributing to public discourse and policy debate by their perception that, in a medical model society dominated by professional discourse, their views would be derided or dismissed—hence speaking out would be futile.

The Stigma of Mental Illness

Foucault's (1972) discursive analysis of the development of the idea of mental illness reveals roots in fear of the unknown and in concepts of sin and moral weakness. Though the new theology is that of science and the medical model supposedly separates mental illness from responsibility and blame (Dain, 1989), strong vestiges of mental illness as moral weakness and of diagnosed patients as persons to be feared, remain. Despite the public "education" campaigns, the absolving promises of biological theories, and the public's permanent interest in faddish diseases (Shorter, 1992), mental illness remains something to be ashamed of and concealed (Link, Mirotznick, & Cullen, 1991). Clients or ex-clients may not wish to expose themselves via public political action (Caras, 1994). If they have a self-blaming victim rather than a survivor orientation, they could share the negative stereotypes associated with the condition and refuse to associate with other clients in group action. (This problem bears an interesting similarity to that faced by victims of sexual harassment/abuse.) The stigma of diagnosed mental illness impose on many persons a form of self-censorship and thus impede their full political participation.

In order to avoid the judgment of moral weakness or poor character clients have an incentive to attribute a morally neutral cause for their symptoms. It is common to hear clients strongly insist that *their* mental illness comes from biochemical causes only. In order to avoid censure under the moral weakness paradigm, clients take refuge in the medical model. The social model, which distributes responsibility among many actors and society as a whole, meets resistance from taxpayers, business, government, established institutions, etc. The family dynamics model has been strongly resisted by family members. More recent "abuse victim" models appear as modified versions of the former model. They have less replaced the medical model than extended its pathologizing aspects to an ever-growing portion of the population. Somatic causes remain the untouched paradigm of choice for those who suffer the most but possess the least power to hold others responsible. Even as its parent science paradigm mutates and we move into the post-industrial/post-modern era, the power of the medical model remains virtually undisputed (Barney, 1994; D. Cohen, 1994).

Other paradigms, including the medical model, will replace the moral weakness paradigm. This is a slow process, however, as the general public has less incentive to reject it than do doctors, families, and clients. Of course, other paradigms may introduce new forms of stigmatization, more or less different from that of moral weakness. Thus, models stressing personal responsibility can result in the stigma of weakness of will. Furthermore, in a society that seems bent on rendering as outmoded social manifestations of kindness and gentleness, the powerless may be disrespected not only because of the deviances that made them powerless, but merely because they are "losers" rather than "winners." Therefore, insofar as the direct or indirect effects of marginalization can be seen as contributing to mental illness and its consequence, loss of power, we can imagine a self-reinforcing dynamic in which a competitive society reproduces insanity as part of a vain and ultimately hopeless attempt of the haves to obtain satisfaction merely by virtue of being haves.

Power Disadvantages Due to Psychological Distress and its Treatments

Distress. Certainly, severe emotional problems must be expected to impair the ability of some individuals to formulate their demands[8] and apply organized pressure to achieve them. It is important to recognize this, being careful not to underestimate the potential of clients to act politically (Rousseau, 1993). Psychological distress does not correlate with low intelligence, and most emotional problems manifest episodically. However, to the extent that those we call the mentally ill are unable to represent their own interests due to their psychological problems, a democratic society should find supportive or alternate ways to give them a voice. A *parens patriae* response is insufficient to address this source of power disability, since much of what is due to distress may in fact have common causes with the distress, which causes should be attacked directly. Poverty, environmental toxicity, unsupportive relationships, and social marginalization due to racism, sexism, physical disability, to the extent that they contribute to psychological distress, can be tackled by education, employment, equality, and other governmental/societal actions (Mirowsky & Ross, 1989b).

Iatrogenesis. Treatment-induced intellectual/social dysfunctions could have more serious impacts on the ability of clients to advance their interests than does the original disorder (Cohen & Cailloux-Cohen, 1995; see Kane & Lieberman, 1992; Keshavan & Kennedy, 1992; Van Putten & Marder, 1987). Breggin (1991, p. 58), who asserts that "all of the major psychiatric treatments exert their primary or intended effect by disabling normal brain function," has coined the term *iatrogenic helplessness* to denote the combination of brain damage and authoritarian suggestion by practitioners administering major psy-

---

[8] However, Appelbaum and Schwartz (1992, p. 463) note that "no systematic body of evidence exists" to support the view that clients suffering severe mental disorders are more "error-prone" in making treatment decisions than the doctors and reviewers of substituted decisions.

choactive drugs and shock treatments. If these treatments do indeed render clients less able or willing to give voice to their experiences and perspectives, the public debate over the mental health system would have to be seen as extremely unbalanced. The possibility of psychiatry in effect silencing its opposition is too repugnant for a democracy to ignore, particularly given the well-documented uses of psychiatry in former communist states to repress political dissent. One need not suggest an insidious plot by psychiatrists, nor even that any psychiatrist would use treatment to exert political authority, to be concerned about the convenient advantage both psychological distress and its treatment accord psychiatrists where their interests diverge from those of clients.

The Clients' "Free-Rider" Problem Versus the Psychiatric Monopoly

Voluntary cooperation is difficult to achieve in large groups. Even though the anticipated gains of cooperation might exceed, for each group member, the required contribution, individuals are predicted to not cooperate where contributions come from individuals and where the benefits are public goods (i.e., where members cannot be excluded from sharing in the benefits). From an individual perspective, the choice to cooperate or not depends on the expected net benefits resulting from the individual's marginal contribution. Not contributing means no cost has been exacted, yet the benefits of others' actions are shared by the noncontributing individual (the "free rider"). The phenomenon of free riders results in "market failure": suboptimal distribution of goods and services. The aggregated undercontributions resulting from individual calculations leave the undercontributors worse off than if they had all given more. In a system where groups compete for resources and favourable laws and policies, those groups able to reduce the free-rider problem by enhancing or enforcing internal cooperation thereby achieve an advantage (Mueller, 1989, pp. 13–14; Thaler, 1994, pp. 6–20).

The free-rider problem derives from the prisoner's dilemma (see Rapoport & Chammah, 1965; Watzlawick, 1976) as developed in public choice theory.[9] Various assumptions must be met before a free-rider program can be predicted,[10] but it can be expected that most large voluntary groups will face this problem to some extent. It is likely that the free-rider problem increases as group size increases (Katz, Nitzan, & Rosenberg, 1990; Olson, 1965)[11] and so the policy advantage clients might be expected to enjoy due to their large numbers compared to psychiatrists is reduced.

There are strong professional and financial incentives to join professional associations such as the American Medical Association and medical specialty

---

[9] Mueller (1989, p. 1) defined *public choice* as "the economic study of nonmarket decision making, or simply the application of economics to political science."

[10] These include certain characteristics of a large group that militate against development of trust due to cooperation, the extent to which members can be excluded from enjoyment of the benefits, and relationships between marginal benefits and marginal costs.

[11] While the tendency of individuals to free ride increases with group size, overall contributions may nevertheless increase, and larger groups may benefit from economies of scale. A number of recent articles have detailed theoretical models or empirically tested implications of the free rider problem; e.g., Heckathorn, 1989; Lipford, 1995; Tuomela, 1992.

associations such as the American Psychiatric Association (APA). Membership is required by some employers, hospitals, and insurance companies. According to Morone (1993), American medical associations "controlled licensure, proscribed unacceptable arrangements in health care organization . . . and dominated legislation on every level of government. In many ways, the medical profession's power constituted an exaggerated case of interest group liberalism—government authority wielded by an industry, generally for the benefit of its members" (p. 726). Membership brings access to association services such as continuing education and publications, which, in the case of the APA, are partly subsidized by pharmaceutical company gifts and advertisements (annual reports of the APA: England, 1992; Sabshin, 1992b). Kilbane and Beck (1990), in an empirical analysis of the free rider problem with respect to the optometric profession and its association in the United States, found results that "strongly support the hypothesis that there will be more free riding in larger groups," that "selective incentives enable latent groups to overcome the free rider problem" (p. 185), and that "By assisting the professional association in overcoming the free rider problem, the state has enabled the association to be a more effective force in influencing other legislation" (p. 186). Similarly, physicians, including psychiatrists, have reduced the free rider problem as they are able to induce contributions in what are ostensibly voluntary organizations.[12] This is evident in the case of the APA by the high proportion of self-identified psychiatrists who belong to the association (approximately 82% in 1989; see Dorwart et al., 1992).

The psychiatric monopoly has been sustained by its complex relationship to the state and judiciary: powers of professional accreditation, of involuntary treatment, the privilege to prescribe drugs and to hospitalize, and, generally, its domination of mental health practices in the hospital system. Physicians also have extensive advantages over other mental health professionals in obtaining health insurance reimbursements and use of hospital resources, and they "exert enormous monopoly powers through the courts" where they are regarded as experts in a large number of spheres (Kiesler, 1992, p. 369). These advantages accorded to psychiatrists enhance their ability to encourage membership and contributions. Obviously, client groups have no such inherent advantages.

To summarize the general impact of the free-rider problem on the ability of client groups to organize and exact contributions: there is little incentive for an individual client to donate time and money to the efforts of a relatively large organization when that client perceives her or his own contribution as making little difference and since he or she will share in the benefits of group lobbying in any event. A doctor in a psychiatric association, on the other hand, may vote to increase fees in the knowledge that his or her increased contribution will be matched, not evaded, by virtually all the other members, and the benefit to be shared will increase accordingly. Therefore, the ability of psychiatrists to turn the profits of lobbying into a private collective good, due to their *de facto*

---

[12] There is, of course, some degree of voluntarism in virtually any organization, as some individuals will make more than the minimum contribution because of altruism (Hechter, 1994) or different perspectives on relative expected costs and benefits.

power to exclude members and therefore enforce rules, gives them a major political advantage over clients. Hence, a public policy intervention ostensibly made to protect the interests of clients—by sanctioning and facilitating self-regulatory bodies expected to maintain acceptable professional standards of care—has had the perverse effect of greatly increasing the relative political marginalization of clients.

External Funding Asymmetries

Some groups receive funding or other income from outside their membership, which can be expected to enhance their capacity for effective action (unless the funding compromises the group's original goals and actions). Although some client groups have received funding from government or private sources, the levels have been notoriously low and insecure (Cohen & McCubbin, 1990). Most funding for client advocacy groups has in effect co-opted the voices of clients. Organizations that attract any significant funding, and in fact stimulate fund raising for psychiatric institutions and research, are variants of groups such as Friends of Schizophrenics and National Alliance for the Mentally Ill (NAMI), which are dominated by relatives of clients and by health professionals. The public perceives that those organizations speak for clients, with other groups constituting the radical fringe.

A group might receive funding from a profit-making company that sees the advancement of the group's objectives as enhancing profits. At present, client groups (as opposed to groups advocating for clients) benefit rarely, if at all, from such funding, as success in their objectives would not significantly benefit any corporation. On the other hand, pharmaceutical companies remain a major source of funding for psychiatric associations (Sabshin, 1992a) as well as for advocacy groups such as NAMI (Breggin, 1991). In addition to the large indirect support for medical research and activities, drug company funding ordinarily comes in the form of advertisements in professional journals and conference subsidies. In 1991, the pharmaceutical industry in the United States spent $352 million on advertisements in medical publications (Fletcher & Fletcher, 1992), many of which are owned by the medical associations.

Aside from the disadvantage to client groups of not receiving drug company funding, it may be argued that such funding seriously compromises its recipients (Jones, 1988). At least one recent study has unambiguously documented that drug-company funding can and does influence physicians to increase prescriptions of particular drugs being promoted (Orlowski & Wateska, 1992). Even the American Food and Drug Administration Commissioner acknowledges that nontraditional methods of drug-industry promotion (in the guise of medical or consumer education, market research, etc.) "may have an even greater potential to mislead than traditional advertising techniques" (Kessler, 1992, p. 951).

The possible compromise of psychiatric practice by pharmaceutical companies has become a pressing concern to the profession: in a 1992 editorial in *Biological Psychiatry* Wortis and Stone are blunt:

> . . . professional psychiatric societies such as the American Psychiatric Association, the American College of Neuropsychopharmacol-

> ogy, and our own Society of Biological Psychiatry are becoming increasingly dependent on drug-company support. The APA now derives less than half its operating income from membership dues . . . The overall influence of the industry . . . is to emphasize drug treatment at the expense of other modalities . . . (pp. 847–848).

As in most conflicts of interest, it is difficult to prove changed behaviour in response to or in anticipation of a benefit. This justifies preventive policies to reduce interest conflicts rather than punitive policies reacting to proven cases of abuse. We suggest, in the strongest terms possible, such a preventive orientation—e.g., severe constraints on the promotion of commercial interests in academic and professional venues—to reduce the appearance of conflicts of interest in psychiatry, which is disturbing and reflects upon the credibility and fair-mindedness of the recipients of the largesse.

Conclusion

We have tried to show, first, that there exists a wide divergence between the interests of psychiatry and clients: none of the three major models underpinning society's trust in psychiatrists justify confidence that the interests of psychiatry and its clients converge enough to warrant psychiatrists' speaking and acting for clients in the development of the mental health system and its policies. Second, the distribution of power between psychiatrists and clients is highly unequal: the voices of clients have been co-opted or submerged by those of other groups, particularly organized psychiatry and family-dominated advocacy organizations. Our argument is not based on any particular conception of what the "needs" of clients are—we have not claimed to know what they are, nor, indeed, that they are determinable. However, our point is that the mental health system remains with no good theory to support a proposition that needs will be met, leaving no basis upon which to evaluate the system's success. Therefore, insofar as the "purpose" of this system is to meet client needs, we consider the system to be irrational.

The numbers of clients and their presumed intensity of interest in mental health policy should have guaranteed them a place of importance in the political processes shaping the mental health system. There are several structural reasons why this has not been the case: client passivity due to the medical model therapeutic context; hesitancy to engage in public action due to the enduring stigma of mental illness; incapacities caused by psychological distress as well as by iatrogenic dysfunction; organizational weakness due to the free-rider problem of voluntary client groups compared with the ability of psychiatry to encourage contributions to its lobbying efforts; marked client disadvantages in obtaining external funding.

If we judge one of the positive features of a liberal democracy to be its stability (in that individuals and groups do not need to resort to violence in order to get a fair allocation of society's goods and costs), we need to be watchful. A political system that systematically disadvantages significant segments of society risks alienating them. While such a situation may (temporarily) benefit a small powerful minority, society as a whole will suffer. Goodwin

(1989, p. 47) noted that "over the post-war period the state has consistently sought to recognize greater levels of mental illness in the community." If this is so, what is the reason? Is mental illness one means of explaining away the failures of late capitalist societies? Does it provide a means for nudging marginalized persons toward docility (Fox, 1978)? Hollingsworth's (1992) analysis of the mental health systems of the United States, Germany and France leads to her suggestion that

> . . . the political economy of mental health care seems to be more telling than the specific features and structures of the health care system. . . . What matters is that people with chronic mental illness (and their relatives) are powerless, usually poor, inadequately represented by professionals, and without leverage on their own (p. 922).

If we accept that powerlessness is not just a consequence but also a cause of many deviances, including psychological distress,[13] it becomes apparent that successful intervention requires attention to issues of power, in addition to individualized "treatment," in order to break the vicious circle.

In summary, clients are severely politically disadvantaged, particularly in comparison with psychiatrists, and this situation of extreme imbalance is very unfair from ethics and democracy theory standpoints, and inefficient, if not dangerous, to society as a whole. There is insufficient justification for a general assumption that the mental health system does or will address the needs of its clients; on this standpoint the system as a whole is irrational. Policy options should be examined that improve the responsiveness of the system to client interests at the therapeutic and policy levels.[14] Appropriate and well-targeted policies will require explicit distinction between the abilities of emotionally impaired persons to *recognize* their interests, to *express* them, and to *advance* them. In general, contemporary analysts of the mental health system must shift attention from clients as the *objects* of policy toward clients as the *agents* of policy.

---

[13] Uttaro and Mechanic (1994), in a large national survey of clients whose families belong to NAMI, found that the most important category of unmet needs was "role restoration" (friends, intimate relations, employment, and keeping busy), which was also highly correlated with a measure of consumers' subjective quality of life. In addition, these researchers found a strong relationship between depressive symptoms and unmet role needs, "suggesting that depressive symptoms may either contribute to role difficulties or be a product of them" (p. 373). It is interesting to note the high percentage of the clients responding in this survey who took psychotropic drugs (88%); indeed, among the "perceived needs," "taking medications" was the only one for which more than 37% of respondents were receiving help. Only 5% of respondents wanted more help in this regard, the lowest percentage among all needs.

[14] Such options might include implementing an assisted client advocacy regime; reducing the legal monopoly powers of psychiatrists; prohibiting professional associations from accepting advertisements and other financial support from profit-making organizations that do significant business with the professionals; providing stable public funding for client self-advocacy groups; progressing toward psychosocial/community approaches in order to reduce passivity and the incidence of iatrogenic illnesses; broadening client participation in, and direction of, mental health policies and organizations; disseminating balanced information regarding what is known about mental illness in order to dispel stereotypes and the intellectual monopoly of the medical model; providing public funding for rights challenges in the courts; promoting provider competition and client empowerment by provision of care alternatives, transparent evaluation mechanisms, consumer vouchers, etc.

We have analyzed the mental health sector as being itself a political and economic system, rather than as merely the outcome of therapeutic discourse within the constraints of a given political and economic environment. We conclude, as a consequence of this changed perspective, that in order to achieve a quality mental health system we urgently need to change how mental health policies are produced. According to public choice theorist Buchanan (1995, p. 149): " . . . the objects of social choice are alternative assignments of rights, or alternative rules structures, rather than alternative social states . . . " In other words, a fair result requires a fair game.

References

Annas, G. J., & Densberger, J. E. (1984). Competence to refuse medical treatment: Autonomy vs. paternalism. *Toledo Law Review, 15*, 561-596.

Appelbaum, P. S., & Schwartz, W. F. (1992). Minimizing the social cost of choosing treatment for the involuntarily hospitalized mentally ill patient: A new approach to defining the patient's role. *Connecticut Law Review, 24*, 433-485.

Balter, M. B., & Uhlenhuth, E. H. (1992). Prescribing and use of benzodiazepines: An epidemiological perspective. *Journal of Psychoactive Drugs, 24*, 63-64.

Barney, K. (1994). Limitations of the critique of the medical model. *Journal of Mind and Behavior, 15*, 19-34.

Bertalanffy, L. von (1968). *General systems theory: Foundations, development, applications*. New York: George Braziller.

Black, B. L. (1986). Institutional context and strategy: A framework for the analysis of mental health policy. In W. R. Scott and B. L. Black (Eds.), *The organization of mental health services: Societal and community systems* (pp. 233-276). Beverly Hills: Sage.

Bloche, M. G., & Cournos, F. (1990). Mental health policy for the 1990s: Tinkering in the interstices. *Journal of Health Politics, Policy and Law, 15*, 387-411.

Bowler, T. D. (1981). *General systems thinking: Its scope and applicability*. New York: North Holland.

Breggin, P. R. (1991). *Toxic psychiatry*. New York: St. Martin's.

Brown, P. (1984). The right to refuse treatment and the movement for mental health reform. *Journal of Health Politics, Policy and Law, 9*, 291-313.

Brown, P. (1985). The mental patients' rights movement, and mental health institutional change. In P. Brown (Ed.), *Mental health care and social policy* (pp. 187-212). Boston: Routledge & Kegan Paul.

Brown, P., & Funk, S. C. (1986). Tardive dyskinesia: Professional barriers to the recognition of an iatrogenic disease. *Journal of Health and Social Behavior, 27*, 116-132.

Buchanan, J. M. (1995). Individual rights, emergent social states, and behavioral feasibility. *Rationality and Society, 7*, 141-150.

Buckley, W. (1967). *Sociology and modern systems theory*. Englewood Cliffs, NJ: Prentice-Hall.

Buckman, R., & Sabbagh, K. (1993). *Magic or medicine? An investigation of healing and healers*. Toronto: Key Porter.

Burstow, B. (1992). *Radical feminist therapy: Working in the context of violence*. Newbury Park, CA: Sage.

Butler, N. M., Campion, P. D., & Cox, A. D. (1992). Exploration of doctor and patient agendas in general practice consultations. *Social Science and Medicine, 35*, 1145-1155.

Callahan, J. C. (1984). Liberty, beneficence, and involuntary confinement. *Journal of Medicine and Philosophy, 9*, 261-293.

Caplan, P. J. (1992, June). What should we ask about women and therapy? *Canada's Mental Health*, pp. 25-27.

Caras, S. (1994). Disabled: One more label. *Hospital and Community Psychiatry, 45*, 323-324.

Castel, R., Castel, F., & Lovell, A. (1982). *The psychiatric society*. New York: Columbia University Press.

Chamberlin, J. (1990). The ex-patients' movement: Where we've been and where we're going. *Journal of Mind and Behavior, 11*, 323-336.

Coburn, D. (1993). State authority, medical dominance, and trends in the regulation of the health professions: The Ontario case. *Social Science and Medicine, 37*, 129-138.

Cohen, C. I. (1993). The biomedicalization of psychiatry: A critical overview. *Community Mental Health Journal, 29*, 509–521.

Cohen, D. (Ed.). (1994). Challenging the therapeutic state, part two: Further disquisitions on the mental health system. *Journal of Mind and Behavior, 15*, 1–202.

Cohen, D., & Cailloux-Cohen, S. (1995). *Guide critique des médicaments de l'âme*. Montréal: Éditions de l'Homme.

Cohen, D., & McCubbin, M. (1990). The political economy of tardive dyskinesia: Asymmetries in power and responsibility. *Journal of Mind and Behavior, 11*, 465–488.

Cohen, D., & Remler, M. (1992). Should the right of mental patients to refuse treatment with psychotropic drugs be severely curtailed? A debate. In E. Gambrill & R. Pruger (Eds.), *Controversial issues in social work* (pp. 301-315). Boston: Allyn & Bacon.

Cohen, D., Thomas, G., Dallaire, B., & Morin, P. (1995). *Cure fermée, dangerosité et expertise psychiatrique: Une analyse quantitative et qualitative des décisions de la Commission des affaires sociales du Québec, 1975-1993*. Unpublished manuscript.

Crozier, M., & Friedberg, E. (1977). *L'acteur et le système*. Paris: Éditions du Seuil.

Dain, N. (1989). Critics and dissenters: Reflections on "anti-psychiatry" in the United States. *Journal of the History of the Behavioral Sciences, 25*, 3–25.

Dewan, M. J., & Koss, M. (1989). The clinical impact of the side effects of psychotropic drugs. In S. Fisher & R. Greenberg (Eds.), *The limits of biological treatments for psychological distress* (pp. 189–234). Hillsdale, NJ: Erlbaum.

Diamond, R. J. (1995). Coercion in the community: Issues for mature treatment systems. *New Directions for Mental Health Services, 66*, 3–18.

Donovan, J. L., & Blake, D. R. (1992). Patient non-compliance: Deviance or reasoned decision-making? *Social Science and Medicine, 34*, 507–513.

Dorwart, R. A., Chartock, L. R., Dial, T., Fenton, W., Knesper, D., Koran, L. M., Leaf, P. J., Pincus, H., Smith, R., Weissman, S., & Winkelmayer, R. (1992). A national study of psychiatrists' professional activities. *American Journal of Psychiatry, 149*, 1499–1505.

Engel, G. (1980). The clinical application of the biopsychosocial model. *American Journal of Psychiatry, 137*, 555–564.

England, M. J. (1992). Report of the treasurer. *American Journal of Psychiatry, 149*, 1431–1434.

England, S. L., & Evans, J. (1992). Patients' choices and perceptions after an invitation to participate in treatment decisions. *Social Science and Medicine, 34*, 1217–1225.

Etzioni, A. (1982). The fight against fraud and abuse: Analyzing constituent support. *Journal of Policy Analysis, 2*, 26–38.

Evans, R. G., & Stoddart, G. L. (1990). Producing health, consuming health care. *Social Science and Medicine, 31*, 1347–1363.

Everett, B. (1994). Something is happening: The contemporary consumer and psychiatric survivor movement in historical context. *Journal of Mind and Behavior, 15*, 55–70.

Fisher, D. B. (1994). Health care reform based on an empowerment model of recovery by people with psychiatric disabilities. *Hospital and Community Psychiatry, 45*, 913–915.

Fisher, S., & Greenberg, R. P. (Eds.). (1989). *The limits of biological treatments for psychological distress: Comparisons with psychotherapy and placebo*. Hillsdale, NJ: Erlbaum.

Fletcher, R. H., & Fletcher, S. W. (1992). Pharmaceutical advertisements in medical journals. *Annals of Internal Medicine, 116*, 951–952.

Foucault, M. (1962). *Maladie mentale et psychologie*. Paris: Presses Universitaires de France.

Foucault, M. (1972). *Histoire de la folie à l'âge classique*. Saint-Amand (Cher), France: Gallimard.

Fox, R. W. (1978). *So far disordered in mind*. Berkeley: University of California Press.

Frank, R. H. (1985). *Choosing the right pond: Human behavior and the quest for status*. New York: Oxford University Press.

Frankford, D. M. (1994). Scientism and economism in the regulation of health care. *Journal of Health Politics, Policy and Law, 19*, 773–799.

Gigliotti, G., & Rubin, J. (1991). The right to refuse treatment: An application of the economic principles of decision-making under uncertainty. *International Journal of Law and Psychiatry, 14*, 405–416.

Gleick, J. (1988). *Chaos: Making a new science*. New York: Penguin.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates*. New York: Anchor.

Goodwin, S. (1989). Community care for the mentally ill in England and Wales: Myths, assumptions and reality. *Journal of Social Policy, 18*, 27–52.

Griffith, D. B. (1991). The best interests standard: A comparison of the state's *parens patriae* authority and judicial oversight in best interests determinations for children and incompetent patients. *Issues in Law and Medicine, 7*, 283–338.

Hechter, M. (1994). The role of values in rational choice theory. *Rationality and Society, 6*, 318–333.

Heckathorn, D. D. (1989). Collective action and the second-order free-rider problem. *Rationality and Society, 1*, 78–100.

Hill, D. (1983). *The politics of schizophrenia*. Lanham, MD: University Press.

Hollingsworth, E. J. (1992). Falling through the cracks: Care of the chronically mentally ill in the United States, Germany, and the United Kingdom. *Journal of Health Politics, Policy and Law, 17*, 899–928.

Holstein, J. A. (1993). *Court-ordered insanity: Interpretative practice and involuntary commitment*. New York: Aldine de Gruyter.

Ingleby, D. (1985). Mental health and social order. In S. Cohen & A. Scull (Eds.), *Social control and the state* (pp. 141–188). Oxford: Basil Blackwell.

Isaac, R. J., & Armat, V. C. (1990). *Madness in the streets: How psychiatry and the law abandoned the mentally ill*. New York: Free Press.

Jones, K. (1988). *Experience in mental health: Community care and social policy*. London: Sage.

Kane, J. M., & Lieberman, J. A. (1992). (Eds.). *Adverse effects of psychotropic drugs*. New York: Guilford.

Katz, E., Nitzan, S., & Rosenberg, J. (1990). Rent-seeking for pure public goods. *Public Choice, 65*, 49–60.

Keshavan, M. S., & Kennedy, J. S. (1992). (Eds.). *Drug-induced dysfunction in psychiatry*. New York: Hemisphere.

Kessler, D. A. (1992). Addressing the problem of misleading advertising. *Annals of Internal Medicine, 116*, 950–951.

Kiesler, C. A. (1992). U.S. mental health policy: Doomed to fail. *American Psychologist, 47*, 1077–1082.

Kilbane, S. C., & Beck, J. H. (1990). Professional associations and the free rider problem: The case of optometry. *Public Choice, 65*, 181–187.

Kirk, S. A., & Kutchins, H. (1992). *The selling of DSM: The rhetoric of science in psychiatry*. New York: Aldine de Gruyter.

Kupers, T. A. (1993). Psychotherapy, neutrality, and the role of activism. *Community Mental Health Journal, 29*, 523–533.

Lamb, H. R. (1994). Only good news is politically correct. *Hospital and Community Psychiatry, 45*, 517.

Link, B. G., Mirotznick, J., & Cullen, F. T. (1991). The effectiveness of stigma coping orientations: Can negative consequences of mental illness labeling be avoided? *Journal of Health and Social Behavior, 32*, 302–330.

Lipford, J. W. (1995). Group size and the free-rider hypothesis: An examination of new evidence from churches. *Public Choice, 83*, 291–303.

Littlewood, R. (1992). Psychiatric diagnosis and racial bias: Empirical and interpretative approaches. *Social Science and Medicine, 34*, 141–149.

Loring, M., & Powell, B. (1988). Gender, race, and DSM-III: A study of the objectivity of psychiatric diagnostic behavior. *Journal of Health and Social Behavior, 29*, 1–22.

MacIntyre, A. (1973). Ideology, social science, and revolution. *Comparative Politics, 5*, 321–342.

Marmor, T. R., & Gill, K. C. (1989). The political and economic context of mental health care in the United States. *Journal of Health Politics, Policy and Law, 14*, 459–475.

McCubbin, M. (1994). Deinstitutionalization: The illusion of disillusion. *Journal of Mind and Behavior, 15*, 35–53.

McCubbin, M. (1996). *Problems in estimating the best interests of medically incompetent psychiatric patients*. Unpublished manuscript.

Mechanic, D., & Rochefort, D. A. (1990). Deinstitutionalization: An appraisal of reform. *Annual Review of Sociology, 16*, 301–327.

Mercier, C. (1989). Le rôle des facteurs subjectifs dans la fidélité à la médication. *Revue canadienne de psychiatrie, 34*, 662–667.

Miedema, B. (1994). Control or treatment? Experiences of people who have been psychiatrically hospitalized in New Brunswick. *Canadian Journal of Community Mental Health, 13*, 111–122.

Mirowsky, J., & Ross, C. E. (1989a). Psychiatric diagnosis as reified measurement. *Journal of Health and Social Behavior, 30*, 11–25.

Mirowsky, J., & Ross, C. E. (1989b). *Social causes of psychological distress*. New York: Aldine de Gruyter.

Moore, M. (1984). *Law and psychiatry: Rethinking the relationship*. New York: Cambridge University Press.

Morone, J. A. (1993). The health care bureaucracy: Small changes, big consequences. *Journal of Health Politics, Policy and Law, 18*, 723–745.

Mosher, L. R., & Burti, L. (1989). *Community mental health: Principles and practice*. New York: Norton.

Mueller, D. C. (1989). *Public choice II*. Cambridge: Cambridge University Press.

Oaks, D. (1993). Clue: Government & industry & phony "consumer" groups unite. *Dendron, 33/34*, 20.

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Orlowski, J. P., & Wateska, L. (1992). The effects of pharmaceutical firm enticements on physician prescribing patterns. *Chest, 102*, 270–273.

Parsons, T. (1951). Illness and the role of the physician. *American Journal of Orthopsychiatry, 21*, 452–460.

Prior, L. (1991). Community versus hospital care: The crisis in psychiatric provision. *Social Science and Medicine, 32*, 483–489.

Rapoport, A., & Chammah, A. (1965). *Prisoner's dilemma*. Ann Arbor, MI: University of Michigan Press.

Reed, S. C., & Lewis, D. A. (1991). The negotiation of voluntary admission in Chicago's state mental hospitals. *Journal of Psychiatry & Law, 19*, 137–163.

Reese, H. W. (1991). Contextualism and developmental psychology. *Advances in Child Development and Behavior, 23*, 187–230.

Regan, T. G. (1987). Negotiating an innovative mental health service: The question of limits to negotiated orders. In D. Coburn, C. D'Arcy, G. M. Torrance, & P. New (Eds.), *Health and Canadian society: Sociological perspectives* (pp. 501–514). Markham, Ontario: Fitzhenry & Whiteside.

Rothman, D. J. (1980). *Conscience and convenience: The asylum and its alternatives in progressive America*. Boston: Scott, Foresman.

Rothman, J. (1989). Client self-determination: Untangling the knot. *Social Service Review, 63*, 598–612.

Rousseau, C. (1993). Community empowerment: The alternative resources movement in Quebec. *Community Mental Health Journal, 29*, 535–546.

Sabshin, M. (1992a, March 10). To aid understanding of mental disorders [letter]. *New York Times*, p. A24.

Sabshin, M. (1992b). Report of the medical director. *American Journal of Psychiatry, 149*, 1434–1444.

Scott, W. R. (1985). Systems within systems: The mental health sector. *American Behavioral Scientist, 28*, 601–618.

Scull, A. (1977). *Decarceration*. New York: Prentice Hall.

Shorter, E. (1992). *From paralysis to fatigue: A history of psychosomatic illness in the modern era*. Toronto: Maxwell Macmillan.

Simmons, H. G. (1990). *Unbalanced: Mental health policy in Ontario 1930–1989*. Toronto: Thompson Educational Publishing.

Simon, R. I., & Sadoff, R. L. (1992). *Psychiatric malpractice: Cases and comments for clinicians*. Washington, DC: American Psychiatric Press.

Smith, A. (1924). *An inquiry into the nature and causes of the wealth of nations*. New York: J. M. Dent & Sons.

Stefan, S. (1992). Leaving civil rights to the "experts": From deference to abdication under the professional judgment standard. *Yale Law Journal, 102*, 639–717.

Stein, H. F. (1990). *American medicine as culture*. Boulder: Westview.

Szasz, T. (1991). *Ideology and insanity: Essays on the psychiatric dehumanization of man*. Syracuse: Syracuse University Press.

Szasz, T. (1993). Curing, coercing, and claims-making: A reply to critics. *British Journal of Psychiatry, 162*, 797–800.

Thaler, R. H. (1994). *The winner's curse: Paradoxes and anomalies of economic life*. Princeton, NJ: Princeton University Press.

Torrey, E. F. (1988). *Nowhere to go: The tragic odyssey of the homeless mentally ill*. New York: Harper & Row.

Tuomela, R. (1992). On the structural aspects of collective action and free-riding. *Theory and Decision, 32*, 165–202.

Uttaro, T., & Mechanic, D. (1994). The NAMI consumer survey analysis of unmet needs. *Hospital and Community Psychiatry, 45*, 372–374.

Van Putten, T., & Marder, S. R. (1987). Behavioral toxicity of antipsychotic drugs. *Journal of Clinical Psychiatry, 48* (supplement), 13–19.

Wade, J. C. (1993). Institutional racism: An analysis of the mental health system. *American Journal of Orthopsychiatry, 63*, 536–544.

Wakefield, J. C. (1992). Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder. *Psychological Review, 99*, 232–247.

Watzlawick, P. (1976). *How real is real? Confusion, disinformation, communication*. New York: Random House.

White, D., & Mercier, C. (1991). Coordinating community and public-institutional mental health services: Some unintended consequences. *Social Science and Medicine, 33*, 729–739.

Wolin, S. (1969). Political theory as a vocation. *The American Political Science Review, 63*, 1062–1082.

Wortis, J., & Stone, A. (1992). Editorial: The addiction to drug companies. *Biological Psychiatry, 32*, 847–849.

Zola, I. K. (1978). Medicine as an institution of social control. In J. Ehrenreich (Ed.), *The cultural crisis of modern medicine* (pp. 80–100). New York: Monthly Review Press.

Population Health

A Call for Breadth (Mental Health) and Depth (Psychosocial Theory)

Michael McCubbin

Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec) H3C 3J7, Canada.

[redacted]

Acknowledgments—The genesis of this article benefited from the exchange of ideas and information among the participants at the Workshop on Population Health (CIAR), Kimberly Ontario, held in June 1997. Thanks are extended to Dr. Bernadette Dallaire (GRASP), for her help in editing the manuscript. I also thank Dr. David Cohen (GRASP) and Professor Marc Renaud (President, Social Sciences and Humanities Research Council of Canada), for having shared with me their piquant views, advice and encouragement over the last several years.

18 September 1997

Abstract—This critical review builds upon two recent events affecting the population health field: the political economy critique by Poland *et al.* (1997), and the finding by Marmot *et al.* (1997) that "control" explains the gradient of health with job level in the Whitehall cohort. The article suggests directions that the field could take in developing the theoretical basis for the framework, directions which should integrate, but not be limited to, issues of "control" and political economy.

The directions this paper points to follow from the object of the population health field — the health of populations rather than abstractly isolated individuals; from the basic epistemological underpinnings of the approach — understanding illness requires understanding health; and from the empirical research suggesting the importance of a concept something like "control" for explaining good health and, more generally, of various factors that seem to have as much to do with the *good life* of individuals and collectivities as they do with health.

Incorporation of mental health into the mainstream of population health research would inject some fresh perspectives into the field, and certain areas of psychosocial theory (critical mental health theory; interactionism; Maslow's hierarchy of needs; communitarian theories of society, identity and alienation; political economy of capitalism and globalism; and empowerment theory) hold promise for advancing understanding of the health gradients and particularly what dynamic phenomena underlie the apparent importance of "control". Such a broadened and deepened framework would be more fruitful for both understanding the empirical results we already have and for generating hypotheses which are on the right track in terms of explaining why some people are healthy and others not.

Keywords—population health, physical health, mental health, social theory, critical theory

Population Health

A Call for Breadth (Mental Health) and Depth (Psychosocial Theory)

Introduction

A growing body of research has pointed to "social" factors[1] as important determinants of health and illness (George, 1996). This research has resulted in new ways of looking at health issues, critiquing the medical model and competing with the "health promotion" school. This new approach has become associated with the term "population health", as popularized by the Population Health Division of the Canadian Institute for Advanced Research (CIAR). The members have been in the forefront of developing and communicating their own "population health framework", which has become the reference standard for the population health field (e.g., Evans *et al.*, 1994; Graubard, 1994).

The population health framework provides a way of organizing knowledge in a dynamic model which serves to pose important new directions for research concerning human health and welfare; as such, it might be considered a new paradigm for understanding health. It was established to facilitate research and policy work aimed to improve the health of *populations*, in recognition that different types of problems may necessitate different types and levels of analysis. The population health framework is intended not so much to overthrow the medical model and health promotion approaches, but rather to supplement them and, within the context of the health of *populations*, provide a richer social, economic and historical context for a variety of health and welfare issues, including matters of medicine, health services, and health promotion.

Until now, the work associated with the population health field has been highly descriptive, and the framework for organizing that work remains rudimentary. The main accomplishment of the field in this first decade of its existence has been to establish the importance of social factors, in conjunction with biological and psychological factors, for explaining the health of populations. The field has been successful in reorienting governments and researchers toward this

---

[1] In this article, references to "social factors" and "social theory" refer to all domains of human interaction, including their cultural and economic aspects.

perspective, with important implications for research in the etiology of health and illness and, consequently, for the nature of the policy directions and priorities.

Poland *et al.* (in press) have recently published a critique of the population health framework as developed mostly by CIAR. They criticized it mainly for lack of social theory, avoidance of the political economy dimension — particularly with respect to the implications of world-wide capitalism — and, especially, for underplaying the crucial role of socio-economic differences for health outcomes. Those authors suggested that while CIAR's own research points strongly to inequalities, the CIAR prescriptions tended toward improving wealth and national products. Poland *et al.* then pointed out the danger of the CIAR arguments: that de-emphasizing health *care* because of the importance of social determinants of health, combined with an argument regarding the benefits of health due to wealth, provides governments with an excuse to simply cut health spending in order to reduce government deficits and debts, thereby (according to the predominant neo-liberal ideology) favouring the competitiveness of the nations concerned and their growth in wealth.

Poland *et al.* (in press) stated that while the population health approach has been a valuable development, their aim was to focus on its weaknesses in order to create constructive suggestions for improvement. While those authors seem to have substantially mischaracterized the CIAR approach — an emphasis upon wealth is far from evident in the corpus of CIAR writings — they were right to point out the weakness of the population health field with respect to social theory and, more precisely, political economy.

However, the article by Poland *et al.* chose to focus upon only one characterization of a social determinant, economic inequality, and one area of social theory, the political economy analysis of capitalism. As will be discussed below, while we have observed gradients of health according to economic hierarchies, it is not yet clear why this is so, and what factors interact with wealth and income distribution to create such gradients. Furthermore, there are undoubtedly other social factors independent of economic inequalities which also have important roles for health outcomes.

The purpose of the present article is to build on the positive elements of the population health field and the critique by Poland *et al.*, particularly given another recent event in the population health field, the publication of an article by Marmot *et al.* (1997): they found that adding a "control" variable makes the gradient of health with job level disappear in a population of U.K. civil servants (the Whitehall cohort). The present article suggests directions that the population health field could take in developing the theoretical basis for the framework, directions which should integrate, but not be limited to, issues of "control" and political economy.

This article will first discuss the basic orientation of the population health approach as a new way of looking at issues of health and illness — it is a precondition to the development of coherent theory that the epistemological bases of the approach are clear. The findings and theoretical implications of Marmot *et al.* (1997) will then be discussed; their findings with respect to "control" may mark a watershed in the development of the population health approach from an empirical emphasis to a more balanced science where theory and data interact dialectically — despite the fact that their article contains virtually no discussion of theory.

Finally, two major dimensions in which the CIAR framework would benefit from further breadth and depth will be discussed. First, incorporation of mental health into the mainstream of population health research would inject some fresh perspectives into a field which, after all, is an application of bio-*psycho*-social perspectives of health and illness. Second, certain areas of psychosocial theory, most already well-established outside mainstream health research, hold promise for advancing understanding of the health gradients and particularly what dynamic phenomena underlie the apparent importance of "control".

The Population Health Paradigm: Epistemological Bases

Kuhn (1970) suggested that new paradigms and theories emerge not so much because the old ones were "wrong", but because the old approaches begin hitting limits in their usefulness. Evans and Stoddart (1990), in the seminal paper in the development of CIAR's population health framework, characterized the need for developing new ways of framing our understanding of health as springing from the diminishing marginal returns of the old approaches, and the opportunity cost of passing up the new. The population health school has mounted a demonstration that establishes, at a minimum, the *prima facie* validity of that proposition. The most impelling evidence drawn upon shows slowing in improvements in health outcomes during a postwar period nevertheless marked with rapidly expanding public and private health care spending in western economies. In this era of government cost-cutting, the search for alternative strategies has become intense.

Among the early results of this search has been the provision of a more developed theoretical basis for viewing "public health", which already had a long tradition but which in this century had become largely overshadowed by the medical model. The new theoretical basis, which might be called the *biopsychosocial perspective,* contains a variety of propositions about the determinants of *illness* that, in sum, suggest that most if not all illnesses result from a complex set of interacting factors having not only genetic, infectious, toxic or traumatic origin, but also having independent sources in human psychology and in social environments (Government of Canada, 1986). The proposition that pathological outcomes have antecedents in non-biological as well as biological factors became timely by the increasing priority of the research question: why do some people get sick and others don't, when the biological causes are the same?

Among the first answers to this question was the recognition that individuals make choices that to some extent determine their exposure to biological risk factors that they face. These choices were characterized as contingent upon the "lifestyles" chosen by people: how they pursued the affair of living their lives. Attempts to change people's lifestyles, through researching and publicizing information about risky behaviours, through encouragement, exhortation, and moralizing, as well as through more tangible constraints and incentives such as bans on cigarette advertising, became the central activity of the *health promotion* field. In retrospect, one might characterize this field as focusing on the psychological aspect of the biopsychosocial perspective of health, insofar as attention was placed upon the perceptions and decision processes of individuals.

The health promotion school has achieved some improvements in population health, notably with respect to illnesses related to lack of exercise, poor diet, smoking and drinking. But a recent report on poverty related mortality in the U.S. (Hahn *et al.*, 1996) contains some sobering news for those relying upon health promotion. That study attributed 6.0% of mortality in 1973 to poverty, a figure that hardly fell 18 years later (5.9%). Not only was there virtually no improvement despite the expansion of the U.S. health care system and the launching of major health promotion campaigns in the 1970s and 1980s, but the authors found that the effect of poverty upon mortality must be largely explained by conditions other than commonly recognized behavioural and biological risk factors.

It is not clear that health promotion programs explain the movement toward healthier lifestyles in the western world during the past several decades. A major experiment aimed at testing the efficacy of the health promotion approach showed that subjects hardly improved lifestyles more than controls (who also improved their lifestyles) [the MRFIT trials, described by Syme, 1996]. Even if the health promotion focus is efficacious for major identifiable risk factors like smoking and drinking, however, its application is seen by adherents of the population health approach as quite limited and temporary. Risky behaviours are notoriously difficult to change at the individual level, which the population health approach suggests is due to the importance of wider social factors in structuring individual choices (Renaud, 1993). Without attention to such social factors, not only is it difficult to change risky behaviours but even when such behaviours are changed, nothing prevents new people from engaging in them as the originally targeted people age.

Skolbekken (1995) found that from 1967 to 1991 the percentage of articles with the word "risk" in Medline increased from 0.1% to 5%, with most of that increase from 1986-1991. Skolbekken characterized health research as infected by a "risk epidemic" which brings with it serious methodological problems and policy errors. The author's concluding speculations are quite interesting in the context of the gradient:

> ... [T]he risk epidemic ... is reflecting the socially constructed reality of a particular culture at a particular time in history. In a global and historical context it may be seen as a luxury problem of the richest part of the world. After all, '*moderation in all things—and moderation in that*' requires a freedom of choice that so far has been denied the majority of humans (p. 302).

Probably the most important reason for dissatisfaction with both the medical and health promotion approaches was a large amount of evidence which showed that illness outcomes could only be partially explained by risk factors having known biological impacts, and that poor health declines as one moves up status scales — a gradient which is not fully dependent upon income and access to health services[2]. Much of the new research in the population health field attempts to explain this apparently social phenomenon, thereby turning attention to the last of the terms married to each other in the expression "biopsychosocial". Hence it could be said that the specificity of the research field known as "population health" rests on two mutually supporting foundations: it is aimed at improving the health of aggregates of people, rather than individual persons in particular cases (for which a medical model orientation *might* be appropriate), and that to do so attention must be paid to social factors in addition to biological and psychological factors.

Seen in MacIntyre's (1977) terms as a narrative rather than as a Kuhnian paradigm, if the population health framework is to succeed, it will be because it can explain both what the medical model and health promotion perspective could, but also more. However, the population health framework may also bring us to a paradigmatic change — not only a new way of thinking but one that is fundamentally *different* from the old. That is because of where the inherent logic of the biopsychosocial model *combined* with a population focus leads us: as the search for causality for illness has been displaced from immediate to more distant causes, in terms of time, place, and social structure, we have been seeing increasing indications that for research and policy to make further strides we have to change the starting point of our analysis from effects — e.g., specific diseases — to causes and conditions. This is because different effects do not have different and mutually exclusive sets of causes.

Hence, while biological research is appropriate for biological effects with biological factors, the urgency for *social science* research in health is obvious when we turn attention to causal factors extending beyond the biological, and to the dynamic systems questions that arrive when we try to understand the ways in which all factors interact in human societies. The particular problem facing analysts of *population* health, as opposed to clinical analysis working backward

---

[2] E.g., Mustard C. A. *et al.* (1997) showed mortality as inversely associated with both income and education independently in Manitoba, which has universal health insurance. Adjustment of education by income leaves education with an association with mortality at least as strong as has income. For a recent review of socioeconomic status measures as health determinants see George (1996).

from illness to biological determinants, is related to the fact that the *level* of analysis is different. Both the Evans and Stoddart (1990) framework and George Engel's (1977, 1980) biopsychosocial model view health outcomes and health care via systemic analysis; it is clear that the former implicitly and the latter explicitly are inspired by and consistent with General Systems Theory (GST) [McCubbin, 1997a; McCubbin and Cohen, 1997]. GST is an epistemological method (a formal theory devoid of empirical content) which leads the analyst to think about groupings of individuals and social structures and how they interact (McCubbin, 1997b).

The first principle of GST is that for complex dynamic systems, e.g. systems involving human behaviour, one cannot understand the whole without looking at the parts nor can one understand the parts without looking at the whole (Morin, 1990). This would imply that even if we can "site" health outcomes as manifestations at the individual level, we could not explain such outcomes without situating the individual within impinging systems — social, economic, political.

More precisely, a systemic approach would appear appropriate for trying to understand health outcomes as a function not only of individuals' power, resources, income, etc., but also of the real or perceived amounts of those goods relative to others — e.g., status. In systems terms, health is a *distinctive* rather than aggregate attribute of social systems ("A distinctive attribute is an attribute of the system not possessed by its elements. It is derived from attributes of the interactions among the elements" [Faber and Scheper, 1997]).

We are reaching the limits of what we can understand by epidemiological approaches which try to explain specific diseases by specific risk factors within simplistic linear models that do not allow for dynamic interaction, as noted by a recent editorial in Lancet:

> The techniques used for risk-factor analysis do not identify cause or mechanisms of disease, so the epidemiology of non-communicable or chronic diseases has been likened to a black box, which genetic and molecular epidemiology might open. Appealing as this potential of molecular epidemiology is, the benefits have not been, and are unlikely to be, at the population level (anon., 1997, p. 229).

For a number of years critics of the medical model have decried the pathological focus inherent in that model, characterizing its usefulness as limited. Rather than a system which puts all its resources into at best patching leaky boats and at worst only bailing them out, prevention of ill health requires attention to good health. While philosophers and other social scientists traditionally addressed the question of what is the good life and how it could be obtained, health research rarely dealt with questions of what is good health and how it could be obtained — leave alone what is a good life, which, as will be discussed below, may share many commonalities with what is good health.

While calls mounted for a focus on "health"[3], as reflected in documents like WHO's "Health for All", the established paradigms implicit or explicit in health research impeded the translation of such a focus into research. Those paradigms contained no conception of health other than as the absence of illness. A conception of *health* requires looking at how people actually live their lives in community. Such a concern dovetails well with the focus of population health research: How do social factors and social structures interact with themselves and

---

[3] It is understood in this article that "health", like "illness", is a social construction. Regardless of the reality of phenomena or things to which users of those terms refer, what we choose to associate with those terms is not determined by that reality. Definition and operationalization of the concept of health has been difficult and controversial. However, research into "health" as opposed to "illness" does not necessitate a concrete definition *a priori*: it is normal to concretize and redefine our objects of study as we observe and reflect upon them. Complicating the task of empirical population research in the etiology of health and illness, the same social/cultural/experiential factors affecting health outcomes also influence how people perceive "health" and "illness", what is a health or illness outcome, and the epistemological sense of an intervention. For example, Dallaire (1995[1997]) describes how social representations of health and illness among three groups of hospital workers in different departments are influenced differentially by the nature of the work they do. Lebeer (1997) provides a fascinating account of how non-compliance with cancer treatment itself becomes diagnosed and psychiatrized. While a theory of health with empirical application has to be conscious of the fact that concepts of "health" and "illness" are socially constructed, a starting point is needed. For population health this might be: "Health is that from which illness diverges". This is not the same as "absence of illness", although it looks similar; the former refers to a *process* while the latter to a *state*. Furthermore, this starting point draws attention to the genesis of alternative processes rather than to the consequences of some of them (illness).

other biological and psychological factors such that some people are healthy and others not?

The most important consequence of the shift to the population health framework as a new *paradigm* — perhaps in fact the essence of this paradigm — is an epistemological change: the criterion for evidence is its pertinence not only to *illness* but also to *health.* This shift was evidenced by the seminal work on stress by Selye and subsequently on coping with stress by Antonovsky. As Antonovsky (1987) noted, given that illness is partly attributable to stress, and that stress is endemic to everyday life, what is amazing is not that people get sick but that many people remain healthy!

Socioeconomic Gradients in Health and the "Control" Variable

The work on stress during the 1960s and 1970s helped provide the empirical foundations for the emergence this decade of the population health framework. Subsequent research focused on what kinds of stress are harmful, what are the social conditions that help create such stress, and why some people stay healthy in the face of stress and others don't. Antonovsky (1987) suggested that stress is more harmful when people feel they don't have the means to *control* the situation creating the stress. He then focused on the kinds of *coping skills* that facilitated handling stress in a healthy way.

Surprisingly, the work of the foremost advocates of the population health approach contains few references to Antonovsky's work; most of those advocates are epidemiologists and health policy analysts living in the Atlantic triangle, whereas Antonovsky was an Israeli sociologist. Nevertheless, the empirical emphasis of the health population school brought its adherents to a similar organizing hypothesis, which might roughly be expressed as: "A major factor in bad health is too much or 'bad' stress with which individuals have difficulty coping" (see Cynader, 1994; Cynader and Mustard, 1997; Hertzman and Mustard, 1997). Promising routes for research include, therefore, determining what is bad stress and under what circumstances it is created, how some people cope successfully with stress, and how *groups* of people cope with stress.

The fact that we can model as part of a population health approach the impact of stress experienced by individuals upon individual health outcomes does not imply that stress has no macro-phenomena aspects. Sources, mediation and consequences of stress all have major social components. Indeed, according to Coburn and Eakin (1993), "While viewed by some as inherently individualistic, the concept of stress can also be a powerful tool linking social structures with individual or group physiological response" (p. 89). Therefore, if this line of research is to bear fruit, it would seem that interventions to improve health should be designed for both the individual level (personal coping skills and circumstances) and the collective level (social structures, systems and processes that contribute to creating bad stress and reduce coping ability).

Perhaps the major factor that brought the population health school to a focus on control and coping skills was the need to explain the "socioeconomic gradient in health", which has been the most striking finding of that school: that controlling for all known risk factors and access to health services leaves a marked and continuous positive relationship between social hierarchy level and health (Aïach and Cèbe, 1994). The seminal example of this was provided by Marmot's longitudinal studies of U.K. civil servants (Marmot *et al.*, 1984), which showed cardiac illness decreasing as employees attain higher ranks of the civil service. Not only did several risk factors vary with grade, but controlling for them still left a large gradient. Marmot's subsequent attempts to "explain the unexplained" (Renaud and Bouchard, 1994) resulted in a paper with major significance, published recently, which makes the gradient disappear when controlling for one further variable: the employee's *control* (or at least perception of control) in the work environment (Marmot *et al.*, 1997).

The "control" variable was operationalized by self-assessed questionnaires consisting of 15 items dealing with "decision authority and discretion". The paper unfortunately contains *no* elaboration of the meaning of the variable being measured nor of theory to explain why it is important, other than to mention that control was one variable in the "job-strain model" in earlier research — a two variable model including job demands and control. The job strain model was not supported empirically with respect to job demands, while control was found to vary with cardiovascular risk.

We do not yet know to what extent Marmot's results regarding control are particular to his population of English civil servants in the 1980s, and whether they can or will be replicated by other researchers in a wide variety of social situations. What we do have at this point is a very promising empirical result which remains, nevertheless, grossly undertheorized. Without theory the findings consist of various measurements which are believed to represent an operationalization within particular circumstances of a variable called "control". In 1995, Seeman and Lewis wrote that "there is relatively little that is definitive containing direct measures of the sense of control" (p. 524). The paper by Marmot *et al.* has not changed this situation. A full theory would permit the operationalization of the concept in radically different contexts. One cannot simply map the set of indicators used in a study of civil servants over to situations involving students, congenitally disabled persons, or African tribal villages, for example.

This problem was acknowledged by Marmot *et al.* (1997), leaving a major challenge for researchers: "... how can psychosocial work conditions account for much of the socioeconomic difference in CHD rates, when such differences are also observed in people beyond working age and in housewives classified by their husbands' occupations?" (p. 239). It would seem, then, that there is something hidden in socioeconomic gradients that finds expression in work-place decision latitude, but also in other ways as well. That would suggest that we still have much work ahead in identifying variable "x" which in Marmot's study was largely captured by his operationalization of "control".

Also, we have to bear in mind that a statistical finding that there is no longer a gradient after inclusion of control and risk factors is *not* the same as saying, without qualification, that the gradient is now explained. The latter assertion would imply establishing causality between the explanatory factors and hierarchical position going entirely from the former to the latter; i.e., that employee's positions are determined by their control (or perceived control) over decision-making. If causality can instead run in the other direction, we certainly could not say that we have dismissed an explanation based on position in favour of one based on control but, rather, that lack of control carries risks for illness which to be reduced may require looking at what about the hierarchy reduces control.

Indeed, control is not the only risk factor which varies with position in Marmot's study; so do smoking, physical activity, and hypertension. Again, a concern with the social determinants of health needs to address what role socioeconomic position may have on such risk behaviours — not merely in the statistical sense of correlation, but also in theory, without which explanation would be impossible, and effective intervention unlikely. An explanatory model should consider how social determinants impact health through indirect as well as direct means, but this is rarely done (George, 1996).

Syme (1991) listed the "major findings" of social epidemiology, with respect to disease risk factors, as mobility, social support, type A behaviour, and stressful life events. However,

> Each of these factors has been studied relatively independently of the others and, to my knowledge, no one has seriously attempted a search for commonalities. Seen from the perspective of control, however, it is possible to suggest that all of these factors are simply different facets or manifestations of control or of its absence (p. 19).

These findings have yet to be systematically related within theoretically well-developed models that provide understanding of processes involved. Saunders *et al.* (1996), in a report on priority areas for population health research in Canada (prepared for the Federal/Provincial/Territorial Advisory Committee on Population Health), stated that "strategic directions for original population health research could be improved by a comprehensive initiative to synthesize the available evidence" (p. 1). Accordingly, they identified as one of five priorities study of the mechanisms through which the already observed associations operate.

The most promising frontier for research in the social aspects of health appears to lie in the direction of "control", not only because of the empirical results relating measures of that variable to health outcomes, but also because of the apparent affinity of this concept to other apparent social determinants: social support, status, mobility, sense of coherence, competence, self-esteem, coping skills, learning abilities, insecurity, predictability, frustration, confidence, powerlessness.

But this frontier is forbidding, given the confusing and diverse array of both risk factors and concepts that seem related to control. Therefore, what is needed now is firstly a kind of "archaeology of control". Such an archaeological undertaking would respect the complexity and abstruseness of a concept that likely has many hidden faces, as a Mayan temple contains within it many more. Secondly we need to connect such an archaeological undertaking with theory that goes beyond control: theory that aims to show why socioeconomic gradients in health exist, and how they are expressed in so-called "risk factors" including control.

Mental Health in Population Health Research

Despite the long-standing public health tradition, and occasional early writings by sociologists, historians and anthropologists which began demonstrating the intertwining of health and the organization of society, major and concerted research efforts in the social determinants of health only began in the late 1970s (Frank J. W. and Mustard, 1994), influenced particularly by the provocative writing of Dubos (1969), Illich (1977) and McKeown (1979). On the other hand, for most of the modern history of psychology and psychiatry social factors were seen as either determinative or as strongly interacting with biological factors to create mental illness[4], emotional distress, or, more generally, personality. Although the World Development Report contains evidence that psychiatric disorders constitute a significant proportion of disability (Bland, 1996), the mainstream of health population research has barely touched upon issues of mental health, except that many writers have briefly added mental health outcomes to lists of somatic outcomes that have been related to social factors[5].

---

4. References in this article to "mental illness" pertain to mental health problems as understood within a particular social construction, that which is authorized by the practice of psychiatry.

5. E.g., Mustard C. A. *et al.* (1997) found, with respect to "mental illness and dementia", an association between household income and treatment prevalence for each of four age groups older than 14 (15-29, 30-49, 50-64, 65+). This was in Manitoba, where there is universal health insurance. This would be interesting, for somewhat different reasons, whether the causal direction looked at is the effect of income upon mental illness or of mental illness upon income, particularly since the

There are several possible explanations for this. As knowledge in somatic illness has progressed, acceptance of the pure medical model has declined in the sciences concerned with such illness. The new knowledge that made this possible is research making clearer how *little* we know about the etiology of illnesses: biological factors alone leave much unexplained. But it is ironic that at the same time as the medical model is declining with respect to somatic health, it has become dominant and growing stronger in the field of mental health (Brown, 1990). The field of psychiatry has become mostly biomedical (Thomas *et al.*, 1996), it has gained enormous influence and power within the health and legal systems, and other helping professions — nurses, social workers, psychologists — seem to have adopted a strategy for competing with psychiatry that rests not so much upon different views of mental health but rather upon gaining a larger place in a biomedically oriented psychiatric system.

---

relationship seems to be stronger than for other illness categories crossed by age group: 32% of them showed significant relationships. Although the figures don't separate dementia from mental illness, if we look at the age groups 15-29 and 30-49, where treatment prevalence of dementia can be presumed to be far smaller than for mental illness, both groups show a significant relationship between adjusted income and treatment prevalence for mental illness and dementia, whereas 37% of other categories crossed by those age groups are so associated. These figures seem to suggest a generally stronger association with income for mental illnesses than for somatic illnesses. Note that the dependent variable, treatment prevalence, is only an indicator for actual illness prevalence. We do not know in that study if poorer persons seek psychiatric help more often, if there is an income-related distribution of preferences for psychiatric versus non-medical help for emotional distress, whether the component of psychiatric treatment that is forced or coerced is more common at lower income levels (thereby increasing prevalence at low income levels), or whether mental illness and/or medical treatment for it has a negative impact on income. In addressing the overall results the authors wrote "The significant achievement of [universal health insurance] programs in equalizing access to medical care across socioeconomic groups has not clearly led to an appreciable moderation in the disparities in health status which were part of the rationale for introducing these programs" (p. 12). The disparities in health status with respect to income seem particularly high for mental illness; this would be true whether lower incomes are a risk factor for mental illness or whether mental illness results in lower incomes. Poorer incomes are part of the dynamics of mental illness as experienced by the Manitoba population.

Another reason the field of mental health has been little touched upon in population health research may be that the population health researchers think that the importance of social factors in mental health outcomes is so obvious as to barely need mentioning. Indeed, when mental health is mentioned by these researchers, it is more often because of the interesting fact that the *same* factors result in widely varying outcomes, including not only physical and emotional health problems, but also crime and other consequences considered harmful to society. This led Syme (1996) to suggest that research targeted on diseases classified on the basis of effects rather than causes is severely limited in its ability to understand why some people get sick and others not; such research uses clinical trials which rely upon homogeneous groups without comorbidity; but comorbidity and heterogeneity is common in the real world.

This paper also takes as a given that social factors are of very great, if not determinative, importance for mental health outcomes. Most biologically oriented psychiatrists will admit, if asked, that social factors mediate in the etiology and severity of most, if not all, mental illnesses. Furthermore, while the scientific literature directly challenging the "brain disease" paradigm of modern psychiatry is relatively small, it convincingly shows that those claims rest far more upon speculation rather than upon evidence meeting scientific standards — biological or otherwise (see the collections of articles edited by David Cohen [1990, 1994a], many of which provide extensive reviews of the main themes of this critical literature). This critical literature rejects explanations for mental illness based primarily on somatic phenomena, and tends to view the genesis of emotional distress within a biopsychosocial perspective; the way in which persons are labeled and treated as having mental illnesses is explained in terms of political economy or cultural theories bearing upon the social control of deviant behaviour.

Given the biopsychosocial perspective in the critical mental health theory, it would seem promising to mate that theory with the population health perspective. Such a marriage offers the interesting prospect of, firstly, integrating or linking together some existing critical theory which is highly promising, but lacking integration, into a more applied and empirically oriented population health field, in the hope that the reformulated theory will be elaborated in such a way as to enable its operationalization and testing at *population* levels.

Secondly, it would be very interesting to relate mental health and physical health outcomes in terms of the same population health theories. Not only might we find commonalities and differences that help to explain why some people get somatically ill, others psychologically distressed, others both, and others neither, but such a process could lead to a reconceptualization of emotional distress and its classifications, with important implications for research and intervention.

The Construction of Etiologies of Emotional Distress

Diagnoses of mental illnesses today usually seek their authority in the DSM-IV classifications, put together by consensus in committees of the American Psychiatric Association. These classifications are entirely based upon symptoms. In the psychiatric field symptoms consist of what individuals say and, to a lesser extent, what they do, since there are no biological tests or markers for mental illnesses and their etiologies. DSM classifications are implicitly considered valid to the extent that there is reliability in diagnosis among physicians and that the treatments indicated for those diagnoses — basically, psychoactive drugs — seem to reduce symptoms more for patients having the target diagnoses than for other patients. Among diagnostic categories reliability and treatment targeting vary from poor to reasonably good, but even if there were high diagnostic reliability and treatment targeting the diagnostic system would be validated only to the extent that it is a system understood to bear on *treatment categories* rather than etiology. If, in fact, the etiologies among mental illness are less mutually exclusive than assumed (see Mirowsky, 1990), research based upon diagnostic categories is less likely to be able to identify what dynamic combinations of factors lead to what results, because the way in which results are conceived were defined *a priori* by diagnostic category.

The logic of this argument might be made more apparent by the following very crude heuristic intended to show a possible implication of diagnostic related research based on a DSM validated as above. Assume there are three diagnoses, A, B, and C, and H stands for emotionally healthy. Assume further that there are only four factors involved in the development of mental illnesses: x, a, b, and c. We might imagine a situation where H involves a, b, and c, in widely varying proportions among different persons, but not x. The other diagnoses also involve a, b and c, but A has high levels of a, B high b, and C high c.

The diagnoses of A, B and C were arrived at because differing responses to three different drugs depended upon whether an individual has high a, b or c. Imagine that in fact all the persons with diagnoses have x, but H does not. By focusing research efforts around particular diagnoses we would likely miss factor x, which in this example accounts for why some people suffer psychological distress and others not. Instead investigation would pursue a, b, and c which does not offer to tell us more than we already know: that certain drugs reduce some undesirable symptoms for people distinguished by different levels of a, b and c. The fact that x is involved in all of the diagnosed conditions — and, indeed, may interact with a, b and c to provide different symptoms and drug sensitivities which might not be the case for healthy people with similar levels of a, b and c — will be missed by such research, unless the etiologies of all the diagnosed conditions are compared with H (or, more simply, if a random sample from the combined population of a, b and c are compared with H).

Research under the medical model has never succeeded in tracing back mental illnesses to some of the same factors as for somatic illnesses, with a very few exceptions pertaining to biological factors (e.g., syphilis), which exceptions are notable for then becoming reclassified as to no longer be considered mental illnesses! This has the remarkable implication that as long as a medical model paradigm is retained with respect to mental illness, the interacting biological and psychosocial etiologies of mental illnesses must always remain in a black box. Therefore, we arrive at the suggestion that in order to clarify the role of *social* factors in psychological distress it might be profitable to do so by adopting an approach which has been developed primarily for studying *somatic* distress: i.e., the population health perspective.

One stream of population health research which may spread into mental health areas is that on early childhood development of the brain. Hertzman (1994) reports evidence that early education is a protective factor against dementia and decline in mental function in late life. Cynader (1994) suggests that chronic stress has a deleterious effect on brain function, and where repeated and prolonged, stress hormones cause neurons to die. He asserts that

> ... emerging evidence suggests that those neurons that are *most ready to learn*, those that are within their critical periods, are the *most vulnerable* to the cell death and degeneration associated with chronic stress. ... In addition, the stresses to which we are exposed early in life, during a critical period, may modify our ability to *modulate and control* responses to stressors later on in life (p. 163).

Toward a Population Mental Health Theory

Much or all psychiatric treatment today relies upon invasive treatments upon the brain (drugs, electroshock, lobotomy), treatment which has recently come to be justified within a mostly implicit but nevertheless well-known and widespread "brain disease" or "chemical imbalance" etiology. Aside from the fact that physical abnormalities with respect to mental illnesses have never been established, that etiology remains because the model implies that the *agent* of disease must be biological. However, no replicated evidence of an important causal role for genetic, physical trauma or infectious agents has been found (Jacobs, 1994).

While some invasive treatments might be appropriate to reduce symptoms, none have resulted in cure. Furthermore, although these treatments may reduce psychiatric symptoms, they also affect emotions, cognition and bodily function in ways which are disabling or uncomfortable for many patients (Breggin, 1990; Frank L. R., 1990). There is good reason to believe that what is a "treatment effect" and what is a "side effect" with respect to invasive psychiatric treatments is a purely definitional matter (see Cohen, 1994b) — a matter not left to patient valuation (Cohen and McCubbin, 1990; Susko, 1994).

The research on brain development offers the prospect that if some types of psychological distress in some people could be related to alteration or development of neurological function due to life experience and social factors, prevention or long-lasting amelioration becomes at least theoretically possible. This is in addition to psychosocial intervention based on an etiology of psychological distress explained by varying combinations of how individuals learn to cope with stress and solve problems, along with the nature of the problems as presented or mediated by the social environment[6]. The brain development and psychosocial perspectives are quite consistent within a biopsychosocial model wherein social factors help to create problems and the ways in which people think and act in their environments, and where social factors also might act through neurological alteration[7].

However, a complete population mental health model needs to do more than account for the genesis of emotional distress at the population level. To be useful for improving the mental health of populations it also needs to identify barriers to improving outcomes. The critical theory in the mental health field is concerned with such barriers, identifying cultural, economic and institutional structures and power that impede recognition and intervention with respect to the social factors involved in creating and perpetuating emotional distress (see: Everett, 1994; Farber, 1990; Gergen, 1990; McCubbin, 1994; McCubbin and Cohen, 1996; Scull, 1990; U'Ren, 1997).

---

6. One CIAR contributor, Ellen Corin (1996), has stressed this latter point. She noted that even comparison studies between cultures and nations have tended to avoid cultural and social variables, or to redefine them as individual characteristics. Research conceived in terms of individual characteristics is unable to capture the important role of social, cultural and economic environments in creating stress and problems, and otherwise influencing the ways in which individuals perceive and cope with life circumstances, opportunities and challenges.

7. In fact, a hypothesis of "neurological alteration" is banal in that *all* thoughts, actions and experiences are manifested in brain activity. The hypothesis is useful only insofar as it suggests a biological pathway for the translation of life experience into psychological distress, or insofar as such alteration can be directly "treated". The latter seems unlikely, because the brain activity involved in emotional distress is widespread and not susceptible to "magic bullet" interventions upon the brain (Barney, 1994; Mandell and Selz, 1995).

The Current State Of Population Health Theory

The most integrated attempt at organizing the theoretical basis of population health remains that originally advanced by Evans and Stoddart (1990), which CIAR refers to as a "framework". The framework consists of what is often referred to as the "plumbing diagram" in that article, but as elaborated implicitly or explicitly within the text of that and other articles by writers closely allied to the population health school (most notably, in the collected works edited by Graubard, 1994, and by Evans *et al.*, 1994). The framework reflects the orientation of CIAR researchers valuing both fundamental and applied research: fundamental research being important in order to develop interventions that work.

Although that framework does in fact contain some important theoretical content — mostly implicit — its main function and value is to guide and constrain the further development of theory. It is like a sketch of a car which conveys the main ideas about what a car is and does, but which doesn't provide enough detail in terms of elements and mechanics to enable someone looking at the sketch to build a working version. The model, clearly inspired by earlier work developing the biopsychosocial model as an application of General Systems Theory (Bateson, 1972; von Bertalanffy, 1968; Engel, 1977 and 1980) is straightforward in its principle guidance to future theoretical development of an integrated population health model:

- it must incorporate a wide variety of biological, psychological and social factors;
- it must be complex, in that the factors interact;
- it must be dynamic, in that, for example, the way in which the health service system operates impacts society at large which then provides feedbacks to that system;
- it must model genesis of health outcomes at the population level (which is of course not to say that what happens at the clinical level is unimportant or not pertinent within a population health model);
- it must try to explain the genesis and maintenance of *health* — understanding the genesis of ill-health then becomes a matter of understanding what went wrong;
- the model must be testable and enable the advancement of empirical knowledge;
- the model's usefulness must be, in the end, tied to its potential for generating interventions that could improve the health of populations.

However, the field of population health seems to have been relying on largely inductive approaches to filling in the framework; their work is mostly empirical, heavily characterized by descriptive epidemiological methodology, and little guided by explicit or elaborate theory (Coburn *et al.*, 1995). The discussion of Marmot's paper above provides one important example of both the benefits gained by this empirical bent and the limitations in relying upon it: what is the significance and meaning of what Marmot measured under the variable of "control", not only for his population of civil servants but especially also for other populations including people not in the work force?

In a systematic review of the literature on the social determinants of health, Linda George (1996) concluded that "social factors have been convincingly demonstrated to be strong predictors of health and mortality. That battle is won and need not continue" (p. 248). Her first recommendation was that "research should focus on the causal and temporal processes that underlie the relationships between social factors and health" (p. 247). This cannot be done by descriptive empirical research alone; elaboration of causal and temporal relationships necessitates theoretical model-building. A process in the population health field of testing hypotheses deduced from competing models, hypotheses which would have to be operationalized such that their relationships with their models are explicit, could help the field move ahead. Given the lack of explicit theory, the empirical results that we have today are interesting, but their usefulness is ambiguous. We have a large number of concepts, all of which are likely interdependent, but since the nature of their logical dependencies or actual interactions is untheorized, we are far from creating order in our understanding of population health out of what amounts today to chaos.

Syme (1991) mentioned two approaches that might help to advance research with respect to the "control" variables and related social determinants of health: a typology of control definitions, and a "somewhat different approach":

> ... to think of control as a "sensitizing" concept —a concept that raises consciousness about an issue and that directs thinking along certain lines but that does not provide specific guidance about definitions or assessment methods. .... [A] major priority in this area of work is the initiation of research that specifically compares the usefulness of various definitions and approaches (p. 21).

Health and the "Good Life"

How might we go about such a task? It is obvious that the complexity inherent in the population health framework requires both multidisciplinary research programs, in order to coordinate research in different fields — economics, biology, epidemiology, psychology, medicine, political science, sociology, social work, anthropology, geography, history, to name only a few. But comparing or integrating different definitions, concepts and approaches requires researchers whose methods and interests are *inter*disciplinary: integrating the parts into a coherent whole is an art in itself, and necessitates some vision of the whole. Putting together 1000 piece puzzles can take a lifetime or an eternity when relying on finding pieces that fit with other pieces when a linear mechanical process is used instead of a complexity reducing process; e.g. imagination in constructing, challenging, and reconstructing mental pictures of what the puzzle might represent when assembled. Indeed, the task facing population health theorists is even more daunting than this, since we can never assume that the pieces of the puzzle are elemental and mutually exclusive of the others.

Furthermore, the interdisciplinarity required may be much broader than we usually conceive of with respect to social aspects of health. The control variable, dealing as it does with decision authority and skill discretion, cannot be understood as concretely as, for example, heavy smoking, or even "repressed hostility", in terms of why it is important for health outcomes. While the latter two phenomena may have themselves a complicated etiology, in their direct impacts upon health the former can be easily understood in biological terms and the latter in fairly straightforward psychobiological terms (at least conceivably, in the new science of psychoneuroimmunology). What has not been so obvious to researchers with respect to those two variables, but becomes increasingly obvious with risk factors like "control", is that there is a close connection between those variables, a healthy life, and what for lack of a better term might be called the *good life* — a life that provides satisfaction, stimulation, fulfillment of duties or roles, self-realization, and/or some other qualities.

The population health research and the longer tradition of health sociology has convincingly shown that health outcomes are strongly influenced by how societies structure themselves, how people conduct and experience their lives, and how people interact. What is new in this research is the realization that the influence of these social factors goes beyond issues like access to resources enabling better nutrition, hygiene, medical care and other means of satisfying biological needs, and surpasses the "toxic lifestyle" behaviours which directly threaten human biology through intoxication, infection, trauma, etc. The research on stress and control suggests that *how we experience life in society has its own direct impacts on health independent of the structuring role of society upon resource variables and biological risk behaviour.*

Although such a proposition might have seemed biologically implausible only a few years ago, the field of psychoneuroimmunology has recently shown that emotion, intellect, experience, and body all interact through a variety of neuroindocrinal and perhaps other messaging systems. The extent of such interaction goes well beyond what we had previously understood (e.g. the feeling of pain, volitional movement of muscles). The paradigm of mind-body dualism is not only under attack (Gabbard, 1994; Lyon, 1993), it seems likely to disappear from credible science entirely, as did the "essentialism" paradigm in biology less than a century ago.

However, the loss of this paradigm will not make life easy for researchers who are accustomed to working in well-defined niches aided by advanced specialized knowledge. As we direct some of the attention focused on ill-health outcomes toward the determinants of health we may find the demarcation between "health" (as opposed to illness) and "welfare" (the good life) becoming very fuzzy. Wilkinson (1994) has viewed health as "a general indicator of welfare and the effects of social and economic change" (p. 61). Even though we will likely maintain that health and welfare are logically, etymologically, and empirically different, we remain with an intimate relationship between the two that suggests that *research in the social determinants of health has to concern itself with the "good life"*: health outcomes are mediated and dependent upon how well life is lived. Since the good life is dependent not only upon nourishing the body and avoiding pain, but also a matter of intellectual, emotional, spiritual and social fulfillment, such research will not only have to draw upon the biological, psychological and social sciences, but also upon those fields of knowledge, speculation and expression dealing with the humanities and ethics.

Multilevel Social Theories of the Good Life and How to Get it

Given the above discussions, this section very briefly discusses certain areas of social theory that might be usefully drawn upon for creating a social theory of population health: human interaction theories dealing with frustration and double-bind; Maslow's hierarchy of needs; communitarian theories of society, identity formation, and alienation; the political economy of capitalism and globalism; and empowerment theory. These theoretical areas were selected as meeting the following criteria:

- they have significant implications at the level of populations;
- they bear on what is the good life, how to get it, and how it is frustrated;
- they are multi-level, sensitive to interaction between individuals, groups, and supra-groups;
- they are consistent with a biopsychosocial perspective of health and welfare;
- they seem *prima facie* pertinent to the concept of "control" as a health determinant.

It is not the purpose of this article to establish the usefulness of these theoretical areas for the population health framework, but rather to stimulate interest in the search for appropriate social theory, particularly given the apparent importance of "control" for at least partly explaining social hierarchy gradients in health. As such, the vast areas of research and thought these themes come from are only hinted at; they are presented here merely as a starting point.

Human Interaction Theories: Frustration and Double-bind

Bateson's work, and especially the collection of essays published as *Steps to an Ecology of Mind* in 1972, was highly influential and very unique, as he integrated a variety of theoretical approaches new for his time — game and information theory, General Systems Theory — in a social psychology remarkable for its deep contextualization of choosing human agents in interaction with complex, dynamic social structures and processes. The work of Bateson, along with other psychologists in the interactionism school (notably Watzlawick [1976] and Varela

[1988]) helped develop theories of double-bind and cognitive dissonance which both show how in circumstances of social interaction frustration, confusion, and feelings of powerlessness can arise. Furthermore, they helped develop the systems concept of "emergence" (Polanyi, 1967) by applying it to human problem-solving: an intractable problem might be solved by reframing it or placing it in a larger context.

Information theory (Campbell, 1982; Wilder, 1979) provided one of the sources for "cognitive dissonance", which has become a key concept in social psychology for explaining distress and apparently irrational behaviour. Antonovsky (1987) drew upon this concept for his "sense of coherence" approach to modeling unhealthy versus healthy reactions to stress. The concepts of double bind, cognitive dissonance, and sense of coherence all seem important to explaining why some people stay healthy and others don't, given stress, and it would seem profitable to integrate theory and findings relating to those concepts within a model of how and why different populations cope differently.

If, for example, "control" were to be seen as enabling or facilitating the finding and implementing of "emergent" solutions to frustrating problems creating double-binds, cognitive dissonance and stress, its relationship to both the good life and health outcomes would become more apparent. Furthermore, such an interpretation of control as related to problem-solving and coping skills suggests interventions insofar as such abilities are learned or learnable. To what degree might the strong gradient of good health with higher education (Defo, 1996; Lahelma *et al.*, 1994) be explainable by emergent problem-solving and coping abilities being improved via education?

It also needs to be noted that problem solving and coping skills can be a community as well as individual characteristic. Dissatisfaction with health promotion interventions with high immediate payoffs but which are not sustained has led some analysts recently to develop the concept of "capacity building": the ability of communities to identify and solve problems for themselves (Hawe *et al.*, 1997; see also Jewkes and Murcott, 1996). This reflects the adage by systems evaluation theorist Ackoff (1974), that a poor solution which gets better is better than a good solution which gets worse.

Maslow's Hierarchy of Needs

Richard Wilkinson (1994), commenting upon the dependence of mortality rates in the developed world upon income inequities rather than upon per capita economic growth, suggested that "This represents a transition from the primacy of material constraints to social constraints as the limiting condition on the quality of human life" (p. 61). If this is so, what theory do we have to explain it? Could use of such theory help to model the genesis of health in different groups and populations? The pertinence of Maslow's (1970) well-known hierarchy of needs springs to mind when posing the question "What is the good life". While every individual's answer to that question is unique, at the collective level we are concerned about norms and distributions of preferences and needs. Reference to the hierarchy might help in understanding, for example, why in the study by Marmot *et al.* (1997) of U.K. civil servants control and discretion in decision-making and selection of skills appears to be more important than social support for health outcomes, whereas other studies with other populations have found social support to be highly important.

For example, Dalgard *et al.* (1996) experimented with 3 interventions in a target group of 45-54 year old women with relatively poor social networks, low quality of life and symptoms of psychological distress. All three interventions resulted in significantly improved social network, quality of life, and mental health, compared to the control group, after 3 year and 6 year follow-ups. What is particularly interesting was that the results were similar regardless of the interventions, which were three types of group activity: physical fitness, porcelain painting, pure social activity. Another study by those authors measured the effectiveness of social support groups for elderly persons recently bereaved. Participation in such groups led to a marked improvement in mental health (the tests relied mainly upon anxiety and depression) compared to the control group. Of special interest here was that "those who improved their social network the most also showed the greatest reduction in symptoms" (p. 606).

The studies by Dalgard *et al.* are among the relatively few empirical studies dealing with mental health within a population health perspective. It is pertinent to note that they were developed within a model that shows how some of the key concepts of population health in general are related and, implicitly, how the application of the model can be made contingent upon the life situations of individuals and groups:

> Our approach toward the prevention of psychiatric disorders is based on an etiological model in which stressors, social support, and mental health are the main elements. According to this model, the risk of mental health problems increases when the individual is exposed to stressors. This risk, however, is modified by social support ... which can reduce the negative effects of stressors. The "buffering effect" of social support has been demonstrated in a number of cross-sectional as well as longitudinal studies ... The mechanism behind the buffering effect of social support seems to be mainly a strengthening of the individual's coping abilities, which of course are also dependent on personality factors; not all people need the same degree of social support in difficult life situations (p. 606).

How, then, might we postulate in a broad population health model the working relationships between social support and/or social networks, with stress, coping ability, and (sense of) control? Are both social support and control independent sources of coping ability? Does social support provide a sense or means of control, making control more fundamental? In either case we need to understand why for some groups control seems independently more or less important than social support. One of a number of candidate hypotheses is suggested by the hierarchy of needs: that in a given population, where degree of social support seems to have a minor role in health outcomes, or where it has a negligible role independent of control, in that population adequate social support is already available for virtually all, and that the variable which then becomes most important for explaining variations in coping ability and hence in health becomes a control variable such as decision making authority and skill use discretion. Such a concept of control could be seen as a factor in "self-realization" — a "higher" need than those related to security and affiliation in Maslow's hierarchy. It then becomes pertinent to ask how the concept of self-realization might enlighten what was captured under "control" in the Marmot *et al.* (1997) study.

One very interesting implication of Maslow's theory relates to the pertinence of a "needs" concept for issues of collective health and welfare policy: self-actualizing persons under that theory can no longer be understood in terms of instinctoid need deficits, since those persons are understood to be largely self-directed: they create their own reality and objectives. Hence, insofar as bad stress results from lesser control over events frustrating the objectives or desires of such persons[8], the operationalization of "control" as a health determinant would have to be intimately related to the *object* of control with respect to particular individuals or groups.

For example, for self-actualizing civil servants, control as a health determinant might be correctly defined as "control over decision-making and selection of skills", insofar as this group sees success as power or autonomy within an organization (Sherman and Smith [1984] empirically tested aspects of Maslow's organization theory, showing that external constraints reduce feelings of self-determination). The health of a group of self-actualizing religious mystics might be more responsive to control as "control over states of mind and emotion". Ghandi might have been a self-actualizer who, having reached that stage, could, consistent with Maslow's theory, choose to deprive himself of the "lower" needs of food and sex, feeling frustration and loss of control to the extent that he did not succeed in such deprivation.

For self-actualizing persons, at least, that which motivates and the pursuance of which leads to frustration or satisfaction cannot be identified by seeing what remains after checking off a standardized list of human needs.

Communitarian Theory: Society, Identity, and Alienation

The meaning of society for humanity, and the sources of identity flowing from social and individualistic sources, have been favoured themes of thinkers running back at least to Antiquity. Plato (1955) approached the understanding of the individual good life by reference to the community, whereas modernistic western thinking, reflected in the atomistic epistemology and relativistic ethics work of Hobbes (1983) and, as postulated by social contract theorists (Locke, 1963;

---

[8] This explanation bears similarity to relative deprivation theory. However, this is not a full theory, providing a scientific explanation for human behaviour, unless the individuals or groups modeled are shown in dynamic interaction with others (Faber and Scheper, 1997). For example, behaviour in reaction to frustration will affect others; their feedback to the original actor might affect that actor's choice of object, ability to attain it, sense of frustration, and manifestations of frustration. Hence the variables object, control and frustration are partly dependent upon each other in a dynamic model.

Rawls, 1971; Rousseau, 1992), manifested in liberal democratic capitalism, has viewed society and the State rather as constructs of individuals.

There seems to be little doubt that the ontology of mainstream human sciences over the last century falls firmly into the latter way of thinking: macro-level explanations of human activity tend to rely on aggregating the biology and psychology of individuals. Such reliance leads to analytic errors even if the basic ontology is correct; but if the human individual and human society can only be properly understood in a dialectical relationship of mutual identity formation, as asserted by what has recently come to be called "communitarian theory" (Etzioni, 1995; Fromm, 1947; Loewy, 1993; Taylor, 1989), ignoring this in the human sciences would mean missing the mark with respect to both why individuals and communities act as they do, and how the meaning of the good life is constructed.

The population health field has been struggling to explain why health outcomes seem so importantly dependent upon human factors which cannot all be reduced to the biological — and those human factors cannot be isolated from the collectivity and individuals' perceptions or experiences of their place in the collectivity (see Etzioni, 1988). In the context of a discussion of possible reasons for Japanese longevity (an issue that remains controversial) Evans and Stoddart (1990) suggested the hypothesis that collective self-esteem is related to health outcomes. What is important about such a proposition is the analytical level which it implies: a psychological characteristic which has a collective manifestation or genesis is unlikely to be identified by study only of individuals within that collectivity, since the only measurement referents are observed behavioural norms at the individual level, rather than objective referents independent of those observed (see Ellencweig, 1992). We could not, for example, observe Japanese collective self-esteem without implicit or explicit comparisons with other countries.

Similarly, it may be quite inefficacious to try to intervene at levels below that where the collective factor operates. In Russia, clinic and local community level intervention to reduce problems related to alcoholism might be very severely hampered by the fact of a powerful national cultural norm associating heavy vodka drinking with power, personal efficacy, and masculine virility. This is not to suggest that only national level interventions would be appropriate, but rather that they might be more so — and that more localized interventions would be more effective if they were designed in the knowledge of the collective factors involved (Bell, 1996; Hunter and Riger, 1986). Communitarian theory, concerned with the creation and reproduction of collective norms, values, representations, ideologies and identity, might therefore help explain both health and distress, psychological and somatic, which health research implicitly based upon individualistic ontologies leave under-explained.

The concept of alienation (whether or not so-named) has been extensively developed and provided with rich materialistic, social and epistemological facets, by writers including Hegel (1952), Feuerbach (1957), Marx (1978), Weber (1958), Fromm (1955) and Foucault (1962), along with social scientists and philosophers influenced by them. Insofar as identity is formed in community, it would be very interesting to think about alienation in a population health context, because not only does the concept help to explain various social and economic needs of individuals (e.g., control over the means of production, bearing a striking similarity to Marmot's definition, when we realize that the vast majority of his sample are likely information workers who produce ideas and symbols), but also because of the way concepts of human alienation are necessarily situated in social processes and institutions. Although alienation and anomie may result in the experience of stress at individual levels, those concepts cannot be reduced to or understood by the individual psychological level alone.

Political Economy of Capitalism and Globalization

Poland *et al.* (in press) have very recently critiqued the health population framework, suggesting that it needs much further fleshing out, particularly with social theory, and that its major weakness was a lack of a political economy context for the dynamics being modeled (see also Coburn *et al.*, 1995). Those authors suggested that the population health findings point strongly to social inequalities as the major social factor implicated in poor health (see Wilkinson, 1994, for such an argument, based upon inter-nation comparisons over time), and that therefore attention must be paid to the way such inequalities are created and maintained. The economic organization of society certainly does seem to be a valuable starting point for increasing understanding of lower grade workers' apparent feelings of powerlessness or other dissatisfaction related to the organization of work

(Wilkinson, 1996), which seems, given the results by Marmot *et al.* (1997), to create a gradient in health outcomes by job level.

At a higher level of analysis, some political economy analysis would seem essential in order to understand how social policy pertinent to health is created — in systemic terms, how policy outputs of the health system are transformed in the wider social-economic-political environment and fed back into the system (e.g., Bernstein and Lennard, 1973; Ehrenreich and Ehrenreich, 1971; Fontaine, 1997). As Evans and Stoddart (1990) remarked, resources have continued to be plowed into a "disease management" system, despite declining returns in terms of health outcomes and more attractive alternative investments that aim at prevention by targeting the social determinants of health. Even very poor countries can drastically improve health outcomes by appropriately targeting limited spending to areas including education, public health, economic empowerment of women and subsidized nutrition (Sen, 1993). If indeed researchers, policy makers and the public are agreed that the way in which resources are currently distributed in the health system allows for little further improvement in the health of populations, we need to understand how to effect change in that system, given that it has shown consistent resistance to change (e.g., in moving toward prevention, deinstitutionalization, and community care: McCubbin, 1994; Morrissey *et al.*, 1985).

Finally, it is preposterous to conceive of putting into effect "health for all" at national levels without consideration of the impact of globalization in all areas of life. Everyone is aware of the rapidly accelerating movements of information, capital and labour across national borders, yet a surprisingly small amount of attention has been placed upon the implications of this phenomenon for the structure and nature of societies and governments. Some analysts have noted the influence of international economic competition among governments and corporations in restructuring or destroying the welfare state, but the nature of the change may be much more fundamental: the *loss of nation-state sovereignty,* in that national governments experience diminished capacities to govern. This trend has been documented by the Lisbon Group (Groupe de Lisbonne, 1995); it noted the opportunities and dangers related to globalization, and has been suggesting the need for global social contracts in order to retain some degree of human control over what have become extremely large and impersonal economic forces.

Population health research may have much to gain from the development and application of theory which characterizes the tendencies of a global economy and its major implications for the organization of human activities. With respect to "control" as a health determinant, do larger international corporations, with diffuse or remote policy and management structures, necessarily lead to lower sense of control or self-efficacy for workers? With respect to social inequities as a health determinant, can governments pursue policies of income redistribution, when large corporations and highly paid workers and entrepreneurs can increasingly evade taxes, minimum wages, unions and labour laws by going to a country which tolerates large disparities?

With respect to social determinants of health in general, when might investments in social capital and infrastructure (e.g., pollution control, education, health services, prevention) be defeated by international currency and government bond speculation, and by corporate evasions of taxes or contributions to finance them, or, alternatively, under what conditions can such programs attract investment? Do competing corporations and governments have an incentive to invest in people (see Thurow, 1995)? Rosenberg (1994) noted that "the time-honored distinction between consumption and investment is breaking down" (p. 137), because of a growing realization that "material well-being may have a positive effect on the future capacity of the human agent to produce" (p. 136).

Seriously integrating issues of health with issues of international political economy is extremely difficult, not so much because it couldn't be done, but because there are few today who want to do it. Academics tend to specialize, such that issues of social structure, economics, and health are departmentalized. Policy planners within government departments concerned with health are inhibited for bureaucratic and political reasons from addressing national and international political economy dynamics. Governments feel that the rules of the international economy, and hence how they impact the social determinants of health, are largely out of their hands.

While the major political economy factors structuring the health and welfare systems of nations may seem virtually inaccessible to policy alteration, concern with them is needed in the population health field for two reasons: 1) insofar as major improvements in health — or declines in health — of societies depend largely on interventions whose feasibility are constrained by the political economy factors, even a remote hope of identifying levers that can make a difference would justify such concern; 2) quick adaptation to irresistible global forces requires identifying those forces and their implications with respect to how people live, work, and experience life and work.

Hence a *useful* population health model conceived as an *open system* — i.e., which models health and welfare systems as dynamic in evolving national and international environments — must be informed by theory as to how that environment is changing and what it means for those systems and, hence, the major determinants of health. Recognition of this point may be seeping into the CIAR work, as evidenced by a recent comment by Michael Hayes (1997):

> Global wealth has become increasingly concentrated in the hands of fewer and fewer corporations and individuals. Disparities in income between Canadians are increasing. .... Maintaining a consciousness about the various influences determining who gets what, where, how, why and under what conditions, and actively trying to influence distributive mechanisms to increase public participation and nurture positive identities surely is working in, and for, public health (pp. 6-8).

Empowerment Theory

"Empowerment" is a concept that grew out of theories of alienation, and has two interacting dimensions, one psychological-social and one political-economic. The first pertains to the internal capacities of individuals and groups of individuals to more effectively pursue the good life, due to the development of their personal resources. The second pertains to the external capacities of aggregates of individuals to improve their lives by acting on their environments; this requires enhanced awareness of their places in the environment and the constraints and opportunities facing them. These two dimensions are both logically and empirically inseparable, and are mutually reinforcing, as recognized by the seminal empowerment theorists Berger and Neuhaus (1977) and Rappaport (1987).

While the theory of empowerment grew equally out of praxis and applied critical theory in the 1970s, in the context of the radical social sciences and grassroots activism, social policy analysts have over the past 10 years increasingly used it as an objective for improving the positions of marginalized groups (e.g.: Gutiérrez, 1995; Riessman and Gartner, 1987; Ward and Mullender, 1991). In the community mental health field empowerment of users has become a major, if not the major objective, aiming to reduce dependence, improve self-insight and decision skills, and, in some cases, to improve the capacities of users to act collectively to obtain benefits from the environment — a process which may also have psychological benefits at the individual level (Speer and Hughey, 1996).

In most references to empowerment in the academic literature pertaining to health and illness the emphasis has been upon empowerment as either a clinical or policy prescription strategy. However, relatively little attention has been placed upon the etiology of health and illness implied by an empowerment strategy, even though increasing powerlessness has been shown to modestly predict increased health problems, independent of other known risk factors (Seeman and Lewis, 1995). In the context of the findings of population health research, particularly regarding the importance of control or sense of control for health outcomes, such attention would seem merited for illuminating the results, suggesting interventions based on those results, and for framing further research questions. For example: toward the psychological-individual level, could empowerment theory suggest a way to improve the health of those at lower levels of social hierarchies, if hierarchies are inevitable, or must all solutions be zero sum? Toward the political level, what does empowerment theory suggest regarding the nature of hierarchies of power, the ways in which they operate, and how they could be altered?

Toward the Integration of Psychosocial Theory into the Population Health Framework

The process of fertilizing population health theory from the streams of psychosocial theory discussed above is not linear: one does not simply look at, say, empowerment theory for what it might have to say about population health and then move on to the next. Each theoretical stream and any theory is imperfect in explaining what it ostensibly explains, but may, rather, catch a glimpse of an object which is itself different from theory to theory. Insofar as those objects represent some important dimension of a human reality which cuts through all of them one could at least hope that, with respect to developing theory pertaining to population health, the fertilizing streams could be demonstrated to be consistent or inconsistent with each other. Constructing a population health theory with the framework as a starting point will mean experimentation in a dialectical process as the various theories are continuously spun around in relation to each other.

The process of population health theory construction should progress toward theory which aims to:

- understand why some people stay healthy and others get ill;
- build upon the population health framework, but allowing for fundamental alteration of the framework if such seems justified according to the other principles;
- retain internal consistency among its various propositions and logical links;
- either incorporate psychological/emotional health and illness or justify its exclusion;
- be testable, either directly or through verification of hypotheses deducted from the theory, and be reasonably consistent with those major empirical findings pertaining to population health which seem to be well-supported, at least as reinterpreted with the evolving theory; and
- have enough depth in terms of the lived experiences of individuals and breadth in terms of collective natures and manifestations as to enable the development of policy prescriptions to improve the health of populations made up of persons living in community.

As various fertilizing theories are compared with each other, it would be at least as important to identify inconsistencies as consistencies. Even where the inconsistencies seem removed from issues of health and illness, if they refer to differences among the fundamental propositions of the compared theories, tracing their implications to population health might reveal anomalies in how we explicitly or implicitly think about health issues, thereby pointing the way to gridlock-breaking research thrusts. For example, Marmot *et al.* (1997) found that control or sense of control was much more important than social support for explaining the gradient. Does this imply an inconsistency with communitarian theory which posits humankind as an essentially social animal? Or do we still need to enrich our understanding of sense of control as a social and not just psychological variable?

The theoretical streams above were partly chosen because of some prima facie potential for consistency and pertinence to what we have already observed in the social determinants of health research. If bad stress, as opposed to good stress, hurts the body and mind, perhaps we are helped in understanding what bad stress means psychologically by looking at concepts of double bind and cognitive dissonance. If bad stress comes from feelings of powerlessness or lack of control, empowerment theory would appear pertinent to understanding what creates powerlessness and what to do about it. Perhaps empowerment will mean being able to rise above double binds. Perhaps "alienation" is the phenomenon that empowerment takes aim at: lack of control over an individual's environment, being manipulated and marginalized. If, consistent with communitarianism, individuals create their identity as social animals — differentiated or otherwise — then alienation might render such identity shaky and tentative, creating cognitive dissonance and unhealthy stress. If control of workers' environments and appropriation of their product lies in some anonymous international boardroom — or, for that matter, in an obscure government bureaucratic process — we might expect to find a process of alienation involving a sense of lack of control. And how alienation is experienced — including *from what* a person or collectivity is alienated — might be better understood within hierarchy of needs theory. The critical theory in mental health might help to bridge the gaps between psychological, social, and biological dimensions of human activity and experience, based as it is within a biopsychosocial perspective of emotional distress.

The point here is not that the above speculations lead to what is most important in terms of population health psychosocial theory, but that the theory which is out there does not merely amount to different flavours for different tastes, or different explanations for different objects of study. They have developed out of millennia of reflection upon the question "what is the good life, and how can (I) we get it?". Now that we are faced with the hypothesis that the good life and the healthy life are somehow connected, it makes sense to go to the psychosocial theory bearing on the good life — or how to get it — in order to, at least, weave a useful narrative explaining what makes some people healthy and others not.

Conclusion

The directions this paper points to follow from the object of the population health field — the health of populations rather than abstractly isolated individuals; from the basic epistemological underpinnings of the population health approach — understanding illness requires understanding health; and from the empirical research suggesting the importance of a concept something like "control" for explaining good health and, more generally, of various factors that seem to have as much to do with the *good life* of individuals and collectivities as they do with health.

Many of these factors, including "control", have to be understood both as psychosocial concepts, in terms of the subjective experiences of persons in community, and as social/political-economy concepts, in terms of how and why those experiences are created, structured, and reproduced. In effect, it seems quite reasonable to believe that the "good life" or the lack thereof, understood in both of those dimensions, has important *direct* as well as mediated impacts on both psychological and somatic health. The distinctions between these two manifestations of health may be much less than we have assumed; not because they both have somatic origins, as assumed by most psychiatric practice today, but because both have etiologies rooted in social as well as biological life. While, as discussed above, it is a mistake to try to linearly trace back a line of causality from *specific* diseases toward "the cause(s)", it is also a mistake to start from the ways people live and experience life and society and then ignore non-somatic outcomes.

In order to understand "why some people are healthy and others not" we have to realize how closely intertwined the good life is with a *variety* of outcomes, desirable and undesirable — whether they be happiness, criminal behaviour, emotional distress, or hypertension. While the object of *health* researchers must, of course, be a set of processes or manifestations more narrowly defined than the foregoing, it may well be that *understanding* those phenomena will require a much larger perspective in terms of the good life, the bad life, and their consequences, than health research has been accustomed to.

It is in such a spirit that this article calls for the *broadening* of the population health framework, in terms of phenomena of interest, to incorporate mental health, and the *deepening* of the framework, in terms of explanatory level, to incorporate psychosocial and political economy theory that helps us to understand what the good life is, how it is created or frustrated, and what happens then. Such a broadened and deepened framework would be more fruitful for both *understanding* the empirical results we already have and for generating hypotheses which are on the right track in terms of *explaining* why some people are healthy and others not.

References

Ackoff, R. L. (1974) The systems revolution. *Long Range Planning* **7,** 2-20.

Aïach, P. and Cèbe, D. (1994) Les inégalités sociales de santé. *La Recherche* **25,** 100-109.

anon. (1997) Putting public health back into epidemiology. *Lancet* **350,** 229.

Antonovsky, A. (1987) *Unraveling the Mystery of Health: How People Manage Stress and Stay Well.* Jossey-Bass, San Francisco.

Barney, K. (1994) Limitations of the critique of the medical model. *Journal of Mind and Behavior* **15,** 19-34.

Bateson, G. (1972) *Steps to an Ecology of Mind.* Ballantine, New York.

Bell N. K. (1996) Responsibilities and rights in the promotion of health: differing positions of the individual and the state. *Social Science & Medicine* **43,** 775-782.

Berger, P. L. and Neuhaus, R. J. (1977) *To Empower People: The Role of Mediating Structures in Public Policy.* American Enterprise Institute for Public Policy Research, Washington.

Bernstein, A. and Lennard, H.L. (1973) The American way of drugging: drugs, doctors and junkies. *Society* **May/June,** 14-25.

Bertalanffy, L. von (1968) *General System Theory: Foundations, Development, Applications.* George Braziller, New York.

Bland, R. C. (1996) International health and psychiatry. *Canadian Journal of Psychiatry* **41,** 11-15.

Breggin, P. R. (1990) Brain damage, dementia and persistent cognitive dysfunction associated with neuroleptic drugs: evidence, etiology, implications. *Journal of Mind and Behavior* **11,** 425-463.

Brown, P. (1990) The name game: toward a sociology of diagnosis. *Journal of Mind and Behavior* **11,** 385-406.

Campbell, J. (1982) *Grammatical Man: Information, Entropy, Language, and Life.* Simon & Schuster, New York.

Coburn, D. and Eakin, J. M. (1993) The sociology of health in Canada: first impressions. *Health and Canadian Society / Santé et Société Canadienne* **1,** 83-110.

Coburn, D. and Poland, B., with members of the Critical Social Science and Health Group (1995[1997]) Comment on: "Why are Some People Healthy and Others Not?". *Health and Canadian Society / Santé et Société Canadienne* **3,** 169-176.

Cohen, D. (ed.) (1990) Challenging the Therapeutic State: Critical perspectives on Psychiatry and the Mental Health System. *Journal of Mind and Behavior* **11** (3 and 4).

Cohen, D. (ed.) (1994a) Challenging the Therapeutic State, Part Two: Further Disquisitions on the Mental Health System. *Journal of Mind and Behavior* **15** (1 and 2).

Cohen, D. (1994b) Neuroleptic drug treatment of schizophrenia: the state of the confusion. *Journal of Mind and Behavior* **15**, 139-156.

Cohen, D. and McCubbin, M. (1990) The political economy of tardive dyskinesia: asymmetries in power and responsibility. *Journal of Mind and Behavior* **11**, 465-488.

Corin, E. (1996) La matrice sociale et culturelle de la santé et de la maladie. In *Être ou ne pas être en bonne santé: Biologie et déterminants sociaux de la maladie*, eds R. G. Evans, M. L. Barer and T. R. Marmor, pp. 103-141. Les Presses de l'Université de Montréal, Montréal.

Cynader, M. S. (1994) Mechanisms of brain development and their role in health and well-being. *Daedalus* **123**, 155-165.

Cynader, M. S. and Mustard, J. F. (1997) Brain development, competence and coping skills. *Entropy: Founders' Network Report in Support of The Canadian Institute for Advanced Research* **1**(1), 5-6.

Dalgard, O. S., Sørensen, T., Sandanger, I. and Brevik, J. I. (1996) Psychiatric interventions for prevention of mental disorders: a psychosocial perspective. *International Journal of Technology Assessment in Health Care* **12**, 604-617.

Dallaire, B. (1995[1997]) Définir et reconnaître: représentations de la santé et de la maladie dans trois départements hospitaliers. *Santé et Société Canadienne / Health and Canadian Society* **3**, 71-98.

Defo, B. K. (1996) Areal and socioeconomic differentials in infant and child mortality in Cameroon. *Social Science & Medicine* **42**, 399-420.

Dubos, R. (1969) *Man, Medicine, and Environment*. Mentor, New York.

Ehrenreich, B. and Ehrenreich, J. (1971) *The American Health Empire: Power, Profits, and Politics*. Vintage, New York.

Ellencweig, Y. (1992) *Analysing Health Systems: A Modular Approach*. Oxford Medical Publications, Oxford.

Engel, G. L. (1977) The need for a new medical model: a challenge for biomedicine. *Science* **196** (8 April), 129-136.

Engel, G. L. (1980) The clinical application of the biopsychosocial model. *American Journal of Psychiatry* **137**, 555-564.

Etzioni, A. (1988) *The Moral Dimension: Toward a New Economics*. Free Press, New York.

Etzioni, A. (ed.) (1995) *Rights and the Common Good: The Communitarian Perspective*. St. Martin's Press, New York.

Evans, R. G., Barer, M. L. and Marmor, T. R. (eds) (1994) *Why are Some People Healthy and Others Not? The Determinants of Health of Populations*. Aldine De Gruyter, New York.

Evans, R. G. and Stoddart G. L. (1990) Producing health, consuming health care. *Social Science & Medicine* **31**, 1347-1363.

Everett, B. (1994) Something is happening: the contemporary consumer and psychiatric survivor movement in historical context. *Journal of Mind and Behavior* **15**, 55-70.

Faber, J. and Scheper, W.J. (1997) Interdisciplinary social science: a methodological analysis. *Quality and Quantity* **31**, 37-56.

Farber, S. (1990) Institutional mental health and social control: the ravages of epistemological hubris. *Journal of Mind and Behavior* **11**, 285-299.

Feuerbach, L. (1957) *Essence of Christianity*. Harper, New York.

Fontaine, D. (1997) Reperes pour une modélisation systemique des programmes de santé (Approach for a systemic modelling of health programs). *Social Science & Medicine* **44**, 503-517.

Foucault, M. (1962) *Maladie mentale et psychologie*. Presses Universitaires de France, Paris.

Frank, J. W. and Mustard, J. F. (1994) The determinants of health from a historical perspective. *Daedalus* **123**, 1-19.

Frank, L. R. (1990) Electroshock: death, brain damage, memory loss, and brainwashing. *Journal of Mind and Behavior* **11**, 489-512.

Fromm, E. (1947) *Man for Himself: An Inquiry into the Psychology of Ethics*. Fawcett, Greenwich Conn.

Fromm, E. (1955) *The Sane Society*. Rinehart, New York.

Gabbard, G. O. (1994) Mind and brain in psychiatric treatment. *Bulletin of the Menninger Clinic* **58**, 427-446.

George, L. K. (1996) Social factors and illness. In *Handbook of Aging and the Social Sciences* (4th Ed.), eds R.H. Binstock and L.K. George, pp. 229-252. Academic Press, Toronto.

Gergen, K. J. (1990) Therapeutic professions and the diffusion of deficit. *Journal of Mind and Behavior* **11**, 353-367.

Government of Canada (1986) *Achieving Health for All: A Framework for Health Promotion.*

Graubard, S.R. (ed.) (1994) Health and Wealth. *Daedalus* **123**(4).

Groupe de Lisbonne (1995) *Limites à la compétitivité: Vers un nouveau contrat mondial.* Boréal, Québec.

Gutiérrez, L.M. (1995) Understanding the empowerment process: does consciousness make a difference? *Social Work Research* **19**, 229-237.

Hahn, R. A., Eaker, E. D., Barker, N. D., Teutsch, S. M., Sosniak, W. A. and Krieger, N. (1996) Poverty and death in the United States. *International Journal of Health Services* **26**, 673-690.

Hawe, P., Noort, M., King, L. and Jordens, C. (1997) Multiplying health gains: the critical role of capacity-building within health promotion programs. *Health Policy* **39**, 29-42.

Hayes, M. V. (1997) Work, identity and public participation. Address to the Alberta Public Health Association, 22 May.

Hegel, G. W. F. (1952) *The Philosophy of Right. The Philosophy of History.* Encyclopaedia Britannica, Toronto.

Hertzman, C. (1994) The lifelong impact of childhood experiences: a population health perspective. *Daedalus* **123,** 167-180.

Hertzman, C. and Mustard, J. F. (1997) A healthy early childhood = a healthy adult life. *Entropy: Founders' Network Report in Support of The Canadian Institute for Advanced Research* **1**(1), 3-4.

Hobbes, T. (1983) *Leviathan.* William Collins, Glasgow.

Hunter, A. and Riger, S. (1986) The meaning of community in community mental health. *Journal of Community Psychology* **14,** 55-70.

Illich, I. (1977) *Limits to Medicine. Medical Nemesis: The Expropriation of Health.* Penguin, Markham Ont.

Jacobs, D. H. (1994) Environmental failure-oppression is the only cause of psychopathology. *Journal of Mind and Behavior* **15,** 1-18.

Jewkes, R. and Murcott, A. (1996) Meanings of community. *Social Science & Medicine* **43,** 555-563.

Kuhn, T.S. (1970) *The Structure of Scientific Revolutions* (2nd Ed.). University of Chicago Press, Chicago.

Lahelma, E., Manderbacka, K., Rahkonen, O. and Karisto, A. (1994) Comparisons of inequalities in health: evidence from national surveys in Finland, Norway and Sweden. *Social Science & Medicine* **38,** 517-524.

Lebeer, G. (1997) La violence thérapeutique. *Sciences Sociales et Santé* **15**(2), 69-98.

Locke, J. (1963) *Two Treatises of Government* (Rev. Ed.; intro, notes, P. Laslett). Mentor, Scarborough Ont.

Loewy, E. H. (1993) *Freedom and Community: The Ethics of Interdependence.* State University of New York Press, Albany NY.

Lyon, M. L. (1993) Psychoneuroimmunology: the problem of the situatedness of illness and the conceptualization of healing. *Culture, Medicine and Psychiatry* **17,** 77-97.

MacIntyre, A. (1977) Epistemological crises, dramatic narrative and the philosophy of science. *Monist* **60,** 453-472.

Mandell, A. J. and Selz K. A. (1995) Nonlinear dynamical patterns as personality theory for neurobiology and psychiatry. *Psychiatry* **58,** 371-390.

Marmot, M. G., Bosma, H., Hemingway, H., Brunner, E. and Stansfeld, S. (1997) Contribution of job control and other risk factors to social variations in coronary heart disease incidence. *Lancet* **350**, 235-239.

Marmot, M. G., Shipley, M. J. and Rose, G. (1984) Inequalities in death: specific explanations of a general pattern. *Lancet* **i**, 1003-1006.

Marx, K. (1978) Economic and philosophic manuscripts of 1844. In *The Marx-Engels Reader* (2nd Ed.), ed. R.C. Tucker, pp. 67-125. Norton, New York.

Maslow, A. H. (1970) *Motivation and Personality* (2nd Ed.). Harper and Row, New York.

McCubbin, M. (1994) Deinstitutionalization: the illusion of disillusion. *Journal of Mind and Behavior, 15,* 35-54.

McCubbin, M. (1997a) A review of the influence of General System Theory in the mental health field. Unpublished manuscript.

McCubbin, M. (1997b) General System Theory as a postmodern epistemology for the social sciences. Unpublished manuscript.

McCubbin, M. and Cohen, D. (1996) Extremely unbalanced: interest divergence and power disparities between clients and psychiatry. *International Journal of Law and Psychiatry* **19**, 1-25.

McCubbin, M. and Cohen, D. (1997) Mental health policy from a systemic perspective: orientations for strategic reform. Unpublished manuscript.

McKeown, T. (1979) *The Role of Medicine: Dream, Mirage, or Nemesis?* Princeton University Press, Princeton.

Mirowsky, J. (1990) Subjective boundaries and combinations in psychiatric diagnosis. *Journal of Mind and Behavior* **11**, 407-423.

Morin, E. (1990) *Introduction à la pensée complexe*. ESF Éditeur, Paris.

Morrissey, J.P., Goldman, H.H. and Klerman, L.V. (1985) Cycles of institutional reform. In *Mental Health Care and Social Policy*, ed. P. Brown, pp. 70-98. Routledge & Kegan Paul, Boston.

Mustard, C. A., Derksen, S., Berthelot, J.-M., Wolfson, M. and Roos, L. L. (1997) Age-specific education and income gradients in morbidity and mortality in a Canadian province. *Social Science & Medicine* (in press).

Plato (1955) *The Republic* (trans. H. D. P. Lee). Penguin, Montréal.

Poland, B., Coburn, D., Robertson, A. and Eakin, J. (in press) Wealth, equity and health care: a critique of a "population health" perspective on the determinants of health. *Social Science & Medicine.*

Polanyi, M. (1967) *The Tacit Dimension.* Anchor Books, Garden City, N.Y.

Rappaport, J. (1987) Terms of empowerment/exemplars of prevention: toward a theory for community psychology. *American Journal of Community Psychology* **15**, 121-144.

Rawls, J. (1971) *A Theory of Justice*. Belknap, Cambridge.

Renaud, M. (1993) Social sciences and medicine: Hygeia versus Panakeia? *Health and Canadian Society / Santé et Société Canadienne* **1**, 229-247.

Renaud, M. and Bouchard, L. (1994) Expliquer l'inexpliqué: l'environnement social comme facteur clé de la santé. *Interface* **mars-avril**, 15-25.

Riessman, F. and Gartner, A. (1987) The Surgeon General and the self-help ethos. *Social Policy* **18**(2), 23-25.

Rosenberg, N. (1994) How the developed countries became rich. *Daedalus* **123**, 127-140.

Rousseau, J.-J. (1992) *Du contrat social*. Flammarion, Paris.

Saunders, L. D., Wanke, M. I., Noseworthy, T. W. and Shores, S. J. (1996) *Identification and Assessment of the Current Status of Population Health Research in Canada and Identified Countries.* Federal/Provincial/Territorial Advisory Committee on Population Health, Ottawa.

Scull, A. (1990) Deinstitutionalization: Cycles of despair. *Journal of Mind and Behavior* **11**, 301-311.

Seeman, M. and Lewis, S. (1995) Powerlessness, health and mortality: a longitudinal study of older men and mature women. *Social Science & Medicine* **41**, 517-525.

Sen, A. (1993) The economics of life and death. *Scientific American* **May**, 40-47.

Sherman, J. and Smith, H. (1984) The influence of organizational structure on intrinsic versus extrinsic motivation. *Academy of Management Journal* **27**, 877-885.

Skolbekken, J.-A. (1995) The risk epidemic in medical journals. *Social Science & Medicine* **40**, 291-305.

Speer, P. W. and Hughey, J. (1996) Mechanisms of empowerment: psychological processes for members of power-based community organizations. *Journal of Community & Applied Social Psychology* **6**, 177-187.

Susko, M. A. (1994) Caseness and narrative: contrasting approaches to people who are psychiatrically labeled. *Journal of Mind and Behavior* **15**, 87-112.

Syme, S. L. (1991) Control and health: a personal perspective. *Advances* 7(2), 16-27.

Syme, S. L. (1996) Rethinking disease: where do we go from here? *Annals of Epidemiology* **6**, 463-468.

Taylor, C. (1989) *Sources of the Self: The Making of the Modern Identity*. Harvard University Press, Cambridge.

Thomas, P., Romme, M. and Hamelijnck, J. (1996) Psychiatry and the politics of the underclass. *British Journal of Psychiatry* **169**, 401-404.

Thurow, L. (1995) Communitarian vs. individualistic capitalism. In *Rights and the Common Good: The Communitarian Perspective*, ed. A. Etzioni, pp. 271-282.

U'Ren, R. (1997) Psychiatry and capitalism. *Journal of Mind and Behavior* **18**, 1-27.

Varela, F. J. (1988) Le cercle créatif: esquisses pour une histoire naturelle de la circularité. In *L'invention de la réalité: Contributions au constructivisme* (trans. A.-L. Hacker), ed. P. Watzlawick, pp. 329-345. Seuil, Paris.

Ward, D. and Mullender, A. (1991) Empowerment and oppression: an indissoluble pairing for contemporary social work. *Critical Social Policy* **11**(32), 21-30.

Watzlawick, P. (1976) *How Real is Real? Confusion — Disinformation — Communication*. Random House, New York.

Weber, M. (1958) *The Protestant Ethic and the Spirit of Capitalism* (trans. T. Parsons). Scribner, New York.

Wilder, C. (1979) The Palo Alto group: difficulties and directions of the interactional view for human communication research. *Human Communication Research* **5**, 171-186.

Wilkinson, R. G. (1994) The epidemiological transition: From material scarcity to social disadvantage? *Daedalus* **123**, 61-77.

Wilkinson, R. G. (1996) *Unhealthy Societies*. Routledge, London.

date last revised: 9 December 1997

Short title: Pure best interests

Toward A Pure Best Interests Model of Proxy Decision-making for Incompetent Psychiatric Patients[1]

Michael McCubbin* and David N. Weisstub**

IN PRESS, INTERNATIONAL JOURNAL OF LAW AND PSYCHIATRY, V. 21(1), 1998

*Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec), Canada, H3C [redacted]

**Philippe Pinel Professor of Legal Psychiatry and Biomedical Ethics, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec), Canada, H3C 3J7.

Correspondence should be addressed to: Michael McCubbin, GRASP, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec), Canada, H3C 3J7. Fax: [redacted]

---

[1] We appreciate the thoughtful comments on various drafts from Professor Michael Perlin of New York Law School and David Cohen, Marie-Claude Roberge, Robert Sullivan, and Bernadette Dallaire, all at GRASP, Université de Montréal. The authors would like to thank the generous support of the Université de Montréal and the Québec Réseau d'éthique clinique chez l'humaine (Unité psychiatrie et psychologie).

Toward a Pure Best Interests
Model of Proxy Decision-making for
Incompetent Psychiatric Patients

Introduction: The Debate Over Forced Treatment

This article was generated to explore certain aspects of the larger ethical problem of forced psychiatric treatment. After providing a critical analysis of the theory and practice of how decisions are substituted for patients[2], we will propose a revised model for substituted decision-making, the "pure best interests" model, based solely and explicitly upon autonomy values. This model was developed with psychiatric patients in mind, whereas virtually all previous studies of substituted decision-making, and the models they developed, were based on the extreme situations of persons having profound and permanent incompetence. We believe that the approach developed in this article, built around the more typical yet less publicized type of person labelled incompetent, could provide a useful new direction to help resolve which is, as will be demonstrated here, a general crisis in the theory and practice of proxy decision-making.

---

[2] In this article, references to a "mentally ill" patient or to "mental illness" pertain only to a diagnosis of mental illness. Similarly, "incompetence" pertains only to a legal or quasi-legal judgment of patient incapacity. The term "mental patient" or "patient" may be used, rather than "client" or "consumer", as a label indicating the role of persons subjected to coerced or mandatory treatment; a person in this role does not freely choose services under a consumer model. "Patient as client" is used where the patient is construed as exercising relatively unfettered informed consent (see Miller R. D., 1994). "Proxy" refers here to a formal or de facto substituted decision maker and may include a process incorporating several persons. The proxy process is understood here to include also the validation and application of advance directives.

The unresolved issue of involuntary treatment has given rise to the most heated arguments in the mental health policy field for more than 20 years — despite the interesting fact that the "efficacy of involuntary treatment is especially insufficiently studied" (Vartiainen, Vuorio, Halonen, & Hakola, 1995, p. 166). In the process, the foundations of argument favouring forced treatment have been shaken by a number of pressing issues:

- the validation of psychiatric criteria (Fisher and Greenberg, 1997; Kirk & Kutchins, 1992; Ross & Pam, 1995);
- treatment induced iatrogenic illnesses (Breggin, 1991; Cohen D. & McCubbin, 1990; Keshavan & Kennedy, 1992; Van Putten & Marder, 1987);
- conflicts between the interests of psychiatrists and patients (Brown, 1984; McCubbin & Cohen D., 1996);
- lack of user power in administration and policy (Chamberlin, 1990; Everett, 1994; McCubbin, 1994);
- the tenuous pertinence of a diagnosis of mental illness to incompetence (Cohen D. & Remler, 1992; Grisso & Appelbaum, 1995; Lee & Ganzini, 1994).[3]

---

[3] These concerns, along with the views opposing them more often seen in the public sphere (calls to protect the public and mentally ill persons from themselves), have forced periodic reevaluations of the laws established 20-30 years ago that ostensibly constrain, define and monitor mandatory treatment (Perlin, 1991). Following deinstitutionalization and establishment of psychiatric services in hospitals and clinics, literature, laws and court decisions regarding patient "medical incompetence" have conceptually distinguished the issues of confinement and treatment (Appelbaum, 1992; Appelbaum & Grisso, 1988; Engelhardt & McCullough, 1981). While this distinction suggested that a patient, even involuntarily committed, may have a "right to refuse treatment" (Brooks, 1987; Cichon, 1989; Perlin, 1990), it has also opened the way to forced treatment and other controls on patient liberty "within the community" (Fennell, 1992; Fulop, 1995), which, combined with "moral panic" at the political level over highly publicized tragic incidents, has led to expansion of guardianship, outpatient commitment, and, for example in the U.K., supervision registers (Atkinson, 1996; Cohen A. & Eastman, 1996; Holloway, 1996).

In the face of concerns expressed by family members, psychiatrists, politicians and members of the public revolving around the "need for treatment" and protection of persons perceived as mentally ill, we are far from resolving the ethical question as to whether imposed treatment is an abrogation of individual autonomy and an insult to human dignity, or whether instead (or despite this) it can be justified as the act of a caring community beneficently acting in the interests of someone unable to help himself or herself (Callahan, 1984; Hermann, 1990; Rosenson, 1993; Tancredi & Slaby, 1977).

However, such justification remains largely implicit rather than based upon systematic argument and supporting research. McCubbin & Cohen D. (1996) have questioned the rationality of the mental health system on the grounds that, given power and interest disparities between patients, psychiatrists, and other actors, it is not unsurprising that the system would be unresponsive to patient needs. One hypothesis briefly assessed by those authors was the "best interests model": that it would be reasonable to expect that caregivers are able and willing to assess and satisfy the actual needs of patients. The best interests model is a key foundation of paternalistic systems of care serving vulnerable persons who may have diminished competence (e.g., mental patients, elderly, children, substance abusers).

The Paternalistic Claim: Meeting the "Needs" of Incompetent Patients

While McCubbin and Cohen questioned the incentives to conscientiously ascertain patient "needs", they left largely unexplored the capacity to determine those interests. While in the literature the concepts of the role of patient as client, of informed consent, of competence/incompetence and advance directives have received much recent attention, relatively little scientific effort has been expended in finding ways to ensure the protection and advancement of the interests of persons once they have been declared incompetent. This holds true for incompetent persons in general, but with respect to psychiatric patients, published discussion of substituted decision-making is almost non-existent.

There is a need for a thorough going evaluation of the decision-making methods we are using on behalf of persons designated incompetent, as observed by Dresser and Whitehouse (1994, p. 6):

> ... [The] current subordination of the incompetent patients' experiential interests is unsatisfactory on several scores. The ethicists' and policymakers' near-obsession with defending the competent person's right to control her future treatment has left the best interests standard inadequately developed and subject to widely varied interpretation. As a practical matter, this omission leaves proxy decisionmakers and clinicians with insufficient guidance about how to resolve the vast majority of real cases ...

This quotation, like most of the literature dealing with proxy decision-making, is addressed primarily to situations of incapacity due to unconsciousness or geriatric conditions. The intense ethical and practical puzzles raised by that literature become even more complicated in the case of psychiatric care (Gutheil & Appelbaum, 1980), where many patients are routinely treated with psychoactive drugs, and albeit less often with ECT, despite their expressed opposition.

Psychiatric patients have lagged behind the developments in informed consent affecting medical care in general. According to Gallagher (1996, p. 23):

> ... [T]he broadly drawn right of both competent and incompetent individuals to decisional autonomy has never been regarded as applicable to the case of mentally ill individuals subject to involuntary treatment. Paradoxically, existing law is considerably more deferential to a frankly incompetent person bent on refusing essential life-saving surgery than it is to an

indisputably competent, involuntarily committed person choosing to forego treatment with lithium.

It has become well-known in the literature that many psychiatric treatment decisions are coerced or imposed without legal authority and a determination of incompetence (Diamond, 1995; Lidz & Hoge, 1993; Miedema, 1994; Reed & Lewis, 1990; Winick, 1991). The power disparities between the patient, doctor, family, government and caregivers leave the patient at a distinct disadvantage in the therapeutic situation, and allow inordinate influence for non-patient actors upon the patient's decision (Hill, 1983; Ingleby, 1985; McCubbin & Cohen D., 1996)[4]. Indeed, the relationship between legal status and coercion may be rather tenuous, as shown by the meta-analysis of Monahan *et al.* (1995), which suggested that nearly half of involuntary patients had no idea of their legal status, and that a quarter of involuntary patients would have chosen voluntary admission if offered the choice. In many cases treatment is imposed on patients of limited competence without the protection of a formal proxy decision-making process, because the patient does not object to the imposition of treatment or accedes to pressures to accept it. Competence is rarely questioned when treatment is not refused.

---

4. Patients may be vulnerable to coercion because of their hesitance to incur the hostility of care-givers for non-cooperation or for treading on their professional turf, because of fear of public exposure, and because of the inaccessability of relevant information (Bloche & Cournos, 1990; Donovan & Blake, 1992). Influence may be exerted upon the patient not only by pressure to select a choice, but in the selection and presentation of the information made available to the patient. Janis and Mann (1977, p. 16) note "the tendency of decision makers to be swayed by the form in which information about risks is packaged and presented". This "framing effect" is discussed in Arrow, 1982; Bursztajn, Chanowitz, Gutheil, & Hamm, 1992; Delquié, 1993; Kong, Barnett, Mosteller, & Youtz, 1986; Malloy, Wigton, Meeske, & Tape, 1992; Mazur & Merz, 1993; McNeil, Pauker, Sox Jr. & Tversky, 1982; Redelmeier, Rozin, & Kahneman, 1993; Sutherland et al., 1991; Wang & Johnston, 1995.

It is obvious that such situations of coerced treatment cannot be seen as serving the interests of the patient and public on the grounds of delegation of authority by the patient as client, and hence the use of waivers of informed consent should be treated with healthy suspicion (Wear, 1993), despite Meisel's (1979) perplexing view that they can foster patient self-determination, providing a justifiable exception to the doctrine of informed consent[5]. It is not easy to monitor whether a reasonable effort has been made by a *de facto* proxy decision maker to act according to the patient's interests. Patient power disadvantages that lead to some degree of *de facto* substituted decision-making also serve to conceal the degree to which the patient's interests are protected. For example, it is rare for patient or ex-patient perspectives to find their way, directly or indirectly, into the social science or mental health treatment and administration literature.

It is therefore difficult to justify coerced treatments in the case of a patient who has not been formally assessed as incompetent. However, there is much controversy over the validity, objectivity and reliability of incompetence determinations[6]. For the most part, this article steers clear of that controversy, concentrating instead on how decisions should be made once a psychiatric patient has been formally designated as incompetent in some area of decision-making.

---

[5] Meisel (1979, p. 459) describes his position this way:

> ... [C]ompelling a patient to receive information which he does not wish to receive or to make decisions which he does not whish to make is to fail to respect that patient's dignity. The result is the denial of the right of self-determination, though in this case the consequence of the exercise of the right is that the patient will <u>not</u> participate fully (or at all) in medical decisionmaking. .... The patient remains the ultimate decisionmaker, but the content of his decision is shifted from the decisional level to the metadecisional level ...

It should be borne in mind, however, that the issues of competency and substituted decision-making can be very intertwined in that there is no point in declaring an individual incompetent if no one else can make better decisions than that individual would have made. Promoting autonomy values may mean allowing the incompetent patient to make "wrong" decisions in some instances; furthermore, it is difficult to separate competence with respect to a particular decision from views of the nature of the decision and its importance.

Models of Proxy Decision-making

Professional Judgment

In the United States today, the minimal constitutional standard at the federal level for psychiatric treatment decisions with respect to institutionalized persons is reliance upon professional judgment of practitioners[7]. While the psychiatrist is supposed to do what is "best" for the patient, no guidance is given on how to determine this. Indeed, "no connection need be established between the treatment choice and the individual's preferences" (Gigliotti & Rubin,

---

[6] See: Auerbach & Banja. 1996; Rosenfeld & Turkheimer, 1995a. 1995b; Rosenfeld. Turkheimer, & Gardner, 1992; Winick. 1996.

[7] In 1986 a district court determined that Mr. Charters was incompetent to decide treatment. "The basis for this conclusion was medical testimony suggesting that because medication was the most beneficial course of action the failure to accept such treatment was evidence of medical incompetence" (Gigliotti & Rubin, 1991, p. 407). At the full district appeal court level (U.S. v. Charters, 863 F.2d 302 [4th Cir., 1988]), the distinction made by a panel of that court between irrationality and disagreement between patient and doctor was discarded, and the district court's reliance on "professional judgment" affirmed. The U.S. Supreme Court declined, without giving reasons, to hear an appeal (110 SCT 1317).

It should be noted that the "*Charters* test," as it will be referred to in this article, is in fact a rearticulation of the "Youngberg standard", as it is applied to individuals incompetent to stand trial who wish to refuse medication. See *Youngberg* v. *Romeo*, 457 U.S. 307 (1982), which sets out its "substantial professional judgment" test at 322-23. For a short discussion of this case, see Perlin, 1994, sec. 2.06 at 200-01.

1991, p. 410). Restraint upon the psychiatrist's decision occurs only if it is a substantial departure from professional judgment, standards or practices (which may mean only the usual practice in a certain institution or situation, rather than professional guidelines) [Cichon, 1989; Perlin, 1990; Stefan, 1993]. Furthermore, U.S. federal constitutional law does not require showing incompetence nor even the need for treatment prior to the exercise of professional judgment (Stefan, 1992). Miller R. D. (1994, p. 250) described a U.S. Supreme Court decision (*Washington* v. *Harper*, 110 S. Ct. 1028 [1990]) as follows:

> While continuing to require hearings before mentally disordered prisoners may be involuntarily medicated. it found that legitimate state interests of security and efficiency permit overriding refusals even of prisoners competent to make treatment decisions, as long as professional judgment has been exercised.

This suggests an alarming dual role for psychiatrists, since that law allows the decision to treat or enforce a treatment to be influenced by non-patient interests (see Gaffin, 1996). The *Harper* case, in particular, suggests a conflict of interest for prison psychiatrists involving at least a breach of confidentiality, since even if their medical recommendations were completely independent of institutional considerations, one has to wonder how the prison officials would know about a prisoner's refusal.

A recent American decision followed the *Harper* decision by finding that "allowing plaintiffs freedom to refuse psychotropic drugs ... may infringe on the constitutional rights of others in the Plaintiffs' environment, and there are no more appropriate means of accomplishing the state's interests in safety and control" (*Hightower* v. *Olmstead* [N.D.Ga. Sept 30, 1996], quoted in Stefan, 1996, p. 6). The U.K. *Mental Health Act, 1983* also permits forced treatment of institutionalized persons, regardless of the common law requirement of incompetence. In introducing the measure to the House of Lords, Lord Elton stated that the measure was for the patient's "own good and for the good of the running of the hospital and for the good of other patients" (quoted in Fennell, 1995, p. 319). A fairly new form of treatment coercion in the community is the use of the criminal probation system; treatment may be enforced as a condition of probation or, on the hand, the probation officer may exercise less tolerance of probation violations when the treating psychiatrist reports patient non-compliance (Solomon, Rogers, Draine, & Meyerson, 1995). This form of forced treatment does not require a finding of incompetence and a substituted judgment.

These decisions fail to respect the firm view in the literature — whether medical, legal or ethical — that imposition of treatment requires a prior finding of incompetence. Furthermore, a consensus has formed in the bioethics literature that incompetence is a necessary but not sufficient condition for the imposition of medically recommended treatment; the professional judgment of a medical practitioner regarding treatment does not assure that the incompetent patient's interests will be protected (Griffith, 1991; Hermann, 1990; Robertson, 1985; Strudler, 1988). Nevertheless, this view does not prevail in the usual practice of psychiatric care, since ordinarily the purpose of declaring medical incompetence has been to impose a predetermined medical treatment. Accordingly, the primary evidence used to assess medical incompetence is often, in effect, the patient's disagreement with the psychiatrist's proposed treatment (Appelbaum & Grisso, 1995; Holstein, 1993; Venesy, 1994).

Whereas many jurisdictions now require proxy decision-making to conform with some version of "best interests" or "substituted judgment" standards (developed largely in response to heavily publicized court cases involving end of life issues), in psychiatric care there is little evidence that proxy decision makers actually follow such standards. Indeed, it seems most likely that thus far proxy psychiatric decisions have usually operated as rubber stamps for medical decisions (Cohen D. et al., 1997; Susman, 1994) — but there has been almost no research which casts light on how proxy decisions in psychiatric care are arrived at.

Best Interests Standard

The best interests standard, as it is generally understood, requires an "objective" weighing of costs and benefits of alternatives facing the patient or, under a formulation that recently has become more common, the decision that a reasonable person might make under the same circumstances (Moskowitz, 1996). This standard is poorly understood and variously interpreted, and might incorporate a paternalistic conception of what society, family members, or doctors think is, in general, good for the patient, rather than exploring what actually would serve the interests of a unique patient having particular values in a specific context (Arras, 1988; Griffith, 1991; Payton, 1992; Perlin, 1990).

Rather than being based upon autonomy values, the ethical justifications for the best interests standard have frequently been shown to ensue from the principles of beneficence and nonmaleficence (Erlen, 1995). While a best interests standard recognizes the individual's need for consideration and care, it "fails to accommodate recognition of the self-constitution of the person" (Hermann, 1990, p. 380). A poignant example of this is provided in Stefan's (1996, p. 7) description of a recent U.S. court decision (*In re Boyle*, 674 A.2d 912 [Me., 1996]):

> Holding that even though a state hospital resident had clearly expressed opposition to taking psychotropic medication when she was competent, and that opposition had lasted over a fifteen year period, and even though she had insisted that she would prefer to spend her life in an institution than be forced to take psychotropic medication, the absence of a written advance

> directive and the facts that (not to put too fine a point on it) the court disagreed with the values embodied in her preference to spend her life in an institution rather than take psychotropic medication, and her doctors and family thought she should take the medication, warranted a finding that she could be forced to take the medication.

Degrazia (1995, p. 51) wrote that "the literature on the best interests standard reveals almost no engagement of recent work in value theory". He showed how different value theories can result in very different decisions under the best interests test with respect to neonatal and persistent vegetative state decisions. The vague fluidity of the best interests concept easily lends itself to incorporating the values of persons other than the patient's, such that proxy decisions are contrary to the patient's interests.

For example, Emanuel E. J. and Emanuel L. L. (1992, p. 2069) had viewed with alarm recent tendencies which interpret best interests decision-making by proxies as giving the family "the power to exercise <u>its right over</u> the patient". These tendencies are further evident in an article by Moskowitz several years later (1996, p. 162) in which, noting the philosophical and practical problems in relying upon advance directives, she argued that where little information is available about current patient preferences, the self-determination principle "offers no meaningful moral guidance for end of life care". In focussing upon the "moral authority of families", emphasizing "values of familial love and interdependence" (p. 166), Moskowitz, like many others recently, believes that the balance should shift from autonomy to beneficence in situations where there has been frustration in operationalizing autonomy values.

Far from basing the best interests test on the autonomy value of right to refuse treatment, such interpretations direct the *parens patriae* justification for substituted decisions toward the good of families and society, rather than toward the direction this concept should be taking, of the state empowering vulnerable individuals. It is important to realize the distinction between *parens patriae* conceived as paternalistic beneficence and as a developmental autonomy-based concept: for "even in parental caring the telos is not simply caring but caring in such a manner so that the child becomes an autonomous, independent human being" (Bernstein, 1995, p. 17).

Substituted Judgment Standard

There is no consensus in the literature over the definition of the substituted judgment standard (Cantor, 1996). Variations in understanding and application are so wide that it is questionable whether there is any standard at all. Perhaps the most common definition, especially as implied by recent applications invoking the name of this standard, is the requirement of a proxy to make the decision that the patient, prior to the onset of incompetence, would have wanted to be made. Most writers and courts prefer this standard to the best interests standard where there is strong evidence as to what the patient's views were prior to the onset of incompetence (Hermann, 1990; Veatch, 1995a). However, the validity and usefulness of this approach, as it has been understood, remains extremely problematic (Emanuel E. J. & Emanuel L. L., 1992; Jecker, 1990). Substituted judgment is essentially a rule of procedure rather than an objective, and at least ostensibly requires convincing evidence as to what the patient's wishes were (Beauchamp & Childress, 1989; Gutheil, Bursztajn, Kaplan & Brodsky, 1987).

Unambiguous and applicable evidence about prior preferences is rare, however, particularly with respect to prospective psychiatric treatment decisions. An estimate from 1993 suggested that only 25% of the American population had some form of advance care planning such as an advance directive or proxy appointment (Emanuel L. L. & Emanuel E. J., 1993). In the absence of such evidence, substituted judgment can become a "legal fiction" (Harmon, 1990), replaced by the best interests test, or simply by the preferences of a proxy decision maker (Griffith, 1991) which might partly reflect state interests (Gutheil & Appelbaum, 1983; Ramsey, 1978). Another problem is that under most interpretations it does not allow for alteration of the patient's preferences, even though the patient's situation during incompetence may imply changed interests and hence changed preferences. Also, the possibility of coercion or implicit pressure influencing advance directives and misrepresentations of prior patient wishes raises serious questions of validity and ethics (Areen, 1987; Buchanan & Brock, 1986).

Hierarchies and Hybrids

The literature and law since the 1980s appears to have entrenched a "hierarchy" approach which requires priority to substituted judgment where there is clear evidence of prior preferences (including advance directives or advance proxy appointments) and to best interests otherwise (Degrazia, 1995; Moreno, 1993; Parry, 1987). There have also been sporadic attempts to adapt the substituted judgment standard to provide an intermediate level where there is less than clear and convincing evidence of the patient's preferences, by allowing a decision "based on the person's values, beliefs, and preferences" (Gordon, 1993, p. 45; Kline, 1992), or by constructing preferences for the patient based on widespread values, when there is no evidence that the patient would think otherwise (Cantor, 1996). Such efforts to establish an intermediate level have yet to receive much systematic attention, and the ideals of these approaches are far from being implemented.[8] While such proposals move in the same direction as this article, development,

---

[8] A constructive development is represented by the current Ontario legislation. See *Substitute Decisions Act, 1992*, S.O. 1992, c. 30, ss. 66(3–4), which is reproduced by Tomossy & Weisstub, 1997, at note 118.

implementation and especially evaluation (and hence further refinement) of an ethically coherent substituted decision-making model are hampered by the lack of a unifying principle to such hierarchies.

Brock (1994) located substituted decision-making criteria in a continuum rather than in a hierarchy: the weight of evidence regarding the person's preferences falls as one moves away from advance directives at one extreme, through substituted judgment, and toward exclusive reliance upon best interests criteria. Baergen (1995a, 1995b) modified this continuum; he would reduce the distance between the substituted judgment and the best interests standards, contesting Buchanan and Brock's (1989) view of advance directives as a principle. Rather, Baergen would view an advance directive as playing an evidentiary role.

Baergen's model explicitly emphasizes beneficence over autonomy: the substituted decision would be made which is most consistent with the patient's fundamental goals, preferences and values. The criterion would not be an estimate of what the patient would have decided if temporarily competent. He aims to factor out the effects of depression, denial and other factors which could work against the fulfillment of the patient's goals.

Unfortunately, Baergen did not explain why an autonomy-based value for decisions on behalf of incompetent patients is any less valid than for "normal" patients who would of course also have difficulties in stressful contexts making decisions that optimize their basic values. He wrote (1995b, p. 373) that

---

The UK Law Commission also recommended a hybrid model for decision-making. See The Law Commission. *Mental Incapacity* (London: HMSO, 1995) at paras. 3.24 - 3.28.

> [t]he SJS's [Substituted Judgment Standard] usual emphasis on individual autonomy would have one eschew paternalism in the vast majority of cases. ... I am inclined to regard this as part of an unfortunate moral fad; although it is hardly stylish to say so, I believe that paternalism is justified more frequently than current discussions suggest.

However, particularly in the case of psychiatric care, we would suggest that the "moral fad" for autonomy-based substitute decision-making has already given way to a revamped, more or less disguised paternalism.

Some writers have explored the relationship between varying degrees of competence and the appropriate deference to be placed upon expressed preferences. This has led Baergen (1995a, 1995b), implicitly, and Appelbaum and Schwartz (1992), explicitly, to in effect integrate competence assessment with clinical decision-making. Such proposals are intriguing and inevitable given the contingent nature of competence, but raise slippery slope concerns, since they lend themselves to a great deal of discretion on the part of clinicians and family members to discreetly discount patient preferences they disagree with. Furthermore, they may be irretrievably paternalistic, by constructing a test of competence which is dependent upon the specific decision to be made and perhaps inevitably, in practice, upon making the "right" decision.

Confusion Between Ends and Means

Baergen and many other recent writers reject autonomy, or reduce its priority at some point on the continuum, not because of an inherent weakness in the autonomy principle, but because of difficulties in implementing it, *i.e.*, estimating what the temporarily competent patient would have decided. We find it inconsistent, however, to change the justification and objective of substituted decision-making according to empirical conditions. Rather, our commitment should be a commitment of principle, one upon which we can attempt to unify decision-making, based upon the fundamental value of individual autonomy. Seen as a metaphysical justification, the autonomy paradigm, namely respect for the person *as a person*, should dominate our thinking on all matters pertaining to substitute decision-making. This article endeavours to explore whether autonomy can serve as an effective signpost for an ethically justifiable, yet practical, set of criteria for subsituted decision-making.

In functional terms, given present economic and scientific limitations relevant to substituted decision-making, the autonomy value has even with the best of intentions only achieved limited application. Once an overarching commitment is made, the conceptual map will be redrawn in stages. Autonomy, as an ideal and paradigm, will begin to take on a realistic hue as our knowledge-base about disease, for example, and its relationship to judgment, can be more fully documented. Over time, we should be able to inform actual and potential users about the decision-making processes affecting them such that increasing numbers of persons will become proactive in maximizing the autonomy principle in practice. Equally, decision-makers will evolve to see their own role as "autonomy-maximers" to the fullest extent permitted within functioning systems. If an ideal is not immediately obtainable, and compromises will occur in practice, the value of the paradigm should not be put into question.

The concepts of hierarchy or continuum represent an advance in that they aim to replace paternalistic with autonomy values where feasible. This tendency has been recently pursued by Tomossy and Weisstub (1997), who described a "hybrid" model which emphasizes as much as reasonably possible evidence as to the person's views — even preferences expressed while incompetent. We follow the arguments in that paper to its logical conclusion: toward a model with one clear over-riding objective, that of autonomy values, posing the question of feasibility in terms of how best to advance those values when little is known or knowable about the patient as an individual.

The "Pure Best Interests" Objective

There is a need for a more integrated, intuitively sensible and ethically transparent approach, building upon the best interest and substituted judgment rules to advance to the extent possible patient self-determination and, importantly, to protect the interests of the patient from the interests of others. Roots for such a model can be found in Robertson's (1985) "broader form" of substituted judgment (but allowing for possible change in the underlying preferences of the patient), in Griffith's (1991) call for a "purer" form of best interests, in Brock's (1994) continuum and Tomossy and Weisstub's (1997) hybrid, and especially in Strudler's (1988) call for a judicial ascription of a constitutionally-based right to self-determination, even for persons considered incompetent.

According to Cantor (1996, pp. 1241-1243):

> ... both the best interests standard and the substituted judgment standard ... are fatally flawed. .... In theory it might be possible to reform rather than discard the substituted judgment and best interests jurisprudence. This is especially so since the underlying object—following the course the patient would have wanted—is the same for constructive preference, for most versions of substituted judgment, and for best interests. Yet the current versions of those doctrines are so confused and muddled that creation of a fresh vocabulary seems preferable to reformation.

Cantor has made a valuable contribution to the literature by developing a "constructive preference" approach for attributing preferences for formerly competent patients who never issued intelligible instructions. This provides one of a number of means of substituting or estimating preferences. But what is rather more urgent is an unambiguous statement of an ethically acceptable objective of substitute decision-making — an objective which is fundamental enough to apply to all types of incompetent patients and compared to which means of substituted decision-making — e.g., constructive preferences — must be justified.

Pure best interests is here understood, first, as an objective, rather than as a rule of procedure: to make the decision that the *present* patient would have made if *temporarily* competent. This means that although a patient is deemed incompetent, her/his "underlying" preferences and values provide the only justifiable basis for substituted decisions affecting that person. The concept of "underlying preferences" serves as a myth, or regulatory ideal, aimed at maximizing autonomy and minimizing dignity offending imposition where the patient has impaired autonomy (see Laor, 1984). In respect for the dignity of an individual when we are not able to adequately understand the individual's preferences, values and interests as expressed by that person, we act as if that person has, nevertheless, some form of life plans, values and desires, which we aim as much as possible to respect.

Operationalizing the objective of pure best interests for an incompetent person by conducting an appropriate and careful inquiry into that person's underlying preferences is a way of answering Strudler's (1988) call for recognition of an incompetent person's right of self-determination. Properly done, rather than making an "abstract metaphysical claim", a pure best interests model "creates a tool for protecting the interests of an incompetent" person (p. 357). Degrazia (1995, p. 56) suggested that "The idea of a non-individualized substituted judgment, if not an oxymoron, at least departs from the spirit of the decision-making standard". His observation is correct, in that the substituted judgment has become associated with clear evidence of preferences. However, when there is no evidence of preferences, does that necessarily imply a "non-individualized" judgment? His own analysis of cases shows how, in effect, an unknowable individual can be uniquely contextualized which, in effect, recognizes that individual as a *person*.

Throughout the approximately two decades of history of the development of models for substituted decision-making, a frequently posed question has been whether and how the substituted judgment and best interests standards differ. Both in terms of theory and practice the boundaries of those approaches have been so fuzzy that it is difficult to delineate exactly how the pure best interests objective differs from those approaches. For example, Tomossy & Weisstub (1997, p. 137) described the blurring of the two standards as follows:

> ... both the best-interests and substituted judgment standards rely intrinsically on value judgments made by a person other than the incompetent adult for whom the decision is being made. The former involves an external assessment and balancing of interests, while the latter requires the decision-maker to "stand in the shoes" of the incompetent person. an act that is in fact rooted in fantasy. ... The diffference is purely semantic, and in reality, any decision made for a mentally incompetent person will inevitably rely on a combination of the two patterns of decision-making. [footnotes ommitted]

To put it roughly, however, the pure best interests objective has two main characteristics which are unambiguous in comparison to the established standards:

1) It is an objective, rather than a procedure. The other standards fail to distinguish between what they *aim* to do and how their aims might be operationalized.

2) It is based solely on *autonomy* values for the *present* patient. The best interests standard, as it has generally been interpreted to date, is based primarily on beneficence, while the substituted judgment standard is aimed not at the present patient but, for the most part, upon an earlier version of that patient before having lost competence.

The remainder of this part will justify our view that substituted decision-making should be based on the autonomy of the *present* patient, not on a past or future superimposition of values or preferences.

The Personhood Problem

A fundamental problem for substituted decision-making has been the question of which temporal set of attributed preferences should guide decisions. An individual's preferences and interests change over time due to changes within (body and psyche) and in response to changes without (environmental stimuli). Proxy decisions have been guided by preference attributions based upon the preferences before the illness began (or during the last period of cogency/normality), or upon anticipated preference set after the medical intervention has begun and the patient is again considered competent, better, or cured.

Historical Reconstruction of the Person

Moorhouse and Weisstub (1996, p. 123) wrote that "Arguably, projecting the stated wishes of a previously competent person onto a presently incompetent person is equivalent to imposing the wishes of one person onto a different person". In the context of a discussion on advance directives, those authors noted further that "the important philosophical issues raised by the personhood argument have not been addressed" (p. 121).

It may be that the theoretical underpinning of the advance decision-making model, as we have known it, is related to a perception of the diminished self as a form of property, a receptacle which can only take on meaning when "re-attached" to a previously viable "self". This process of metaphysical incorporation amounts in practical terms to the surrender and denial of the possibility of the reconstitution of the self, however fractured. Therefore, this analogy with advance decision-making, such as in relation to end-of-life or after-life decisions, is warranted. Although advance decision-making may result in the sacrifice of our proposed autonomy-based pure best-interests paradigm, we are not, however, prepared to deny the relevance or utility of advance decision-making where the criteria used and their application to present circumstances do not conflict with the autonomy principle, except in situations where the outcome could be viewed as perverse.

Attributing a property right to a person over the disposal of that person's body and possessions after death is not controversial, but there have been few if any defenders of the explicit thesis that the rights and interests of an incompetent person must be subservient to a property interest of an earlier competent person. Such a view, distinguished from holding that prior expressed preferences should be binding because it is the best practical means of giving autonomy to the present incompetent patient, should be overtly justified if it is to play any role at all in the development of decision-making objectives.

Construction of a decision-making model from the starting point of those psychiatric patients who are designated as incompetent — whose incompetence is more often than not partial and temporary — would seem to show intuitively that for the general model there is no place for such a property right. Justification for the complete ethical inviolability of prior preferences would appear to lie on the right of a person to prospectively dispose of that person's body as desired, after that body has lost its status *as a person*. Unless we wish to assert that an individual loses status as a person upon being designated incompetent, the latter premise is untenable. We are left then, with the appropriate role of expressed prior preferences and values in making decisions on behalf of the current patient, having the objective of making the decision the *current* patient would make if competent.

User groups have recently been very active in promoting the use of advance directives as a means of promoting patient autonomy; such would indeed be an important advance. Users would nevertheless be prudent to also take into consideration how non-patient interests can influence advance directives. They may wish some consideration given to what they say after they have been designated as incompetent and become subject to the interpretation by proxies of their prior directives. Appelbaum and Schwartz (1992, pp. 443-444) warn that accepting that mental illness may cause changes in preferences, thereby providing a case for accepting prior rather than current preferences, "raises [the] possibility of tyranny ... in disregarding deviant preferences on purportedly objective grounds, which may, in fact, reflect majoritarian ideological biases" (pp. 443-444).

An important recent case (*In re Martin*, 538 N.W.2d 399 [Mich., 1995]) described by Dresser (1996) demonstrates poignantly the ethical and evidentiary problems that can be raised by too much reliance upon traces of prior expressions of preferences. In this case a man who had been in a car accident remained at least partially conscious and with very limited voluntary movement, and dependent upon a gastrostemy tube. His wife, acting as proxy, conveyed the patient's prior verbal statements of never wanting to live as a disabled person. She sought to have the tube removed, but was opposed by other family members. One physician had testified, however, that the patient had indicated by head nods that he did not suffer or experience pain, and that he did not have a desire to stop living. Among the issues here are the possibility of a personal interest on the part of the proxy, influencing her interpretation of her husband's wishes, what weight to place on the patient's apparent current attitudes (although it was agreed that he was not competent to make medical decisions), and, given likely changes in the personality, intellectual capabilities, memory and interests of the patient, whether the apparent prior preferences of the patient should outweigh on ethical grounds an also uncertain assessment or construction of his current preferences.[9]

The general tendency of decision makers to overestimate the applicability to current circumstances of information about a prior known instance (Beahrs & Rogers, 1993; Greenberg, Pyszczynski, Warner, & Bralow, 1994) could lead to error in proxy decision-making. Finding that a majority of 150 mentally competent dialysis patients wanted to provide some leeway for overriding their advance directives, Sehgal *et al.* (1992, p. 59) concluded that "Strictly following all advance directives may not truly reflect patients' preferences". Indeed, a

---

[9] For a general discussion of how these issues relate to the the context of euthanasia. see Weisstub. 1997.

recent major study of end of life decisions shows not only that many physicians ignore or misunderstand prior expressions of preference, but also implies that most patients do not want to discuss such decisions in advance (Connors *et al.*, 1995). We do not know to what extent such research would be applicable to advance directives with respect to a very different type of complex issue such as forced care.

Tsevat *et al.* (1995) found that the health utilities of surviving seriously ill patients increased over time. This result, while in opposition to some early studies, is consistent with more recent findings regarding colostomy utilities by Boyd, Sutherland, Heasman, Tritchler and Cummings (1990, p. 66), who raised the hypothesis that "utilities for a particular state of health may change when an individual enters that state". This hypothesis was shown as feasible under certain plausible microeconomic assumptions by Johannesson's (1996) model demonstrating that as health status improves willingness to pay for health interventions may increase. Cohen L. M. *et al.* (1993, p. 399) showed a "substantial difference between the patients' considerations of and actual later decisions to terminate dialysis. ... [C]hronically ill people may have different attitudes regarding cessation of life support than individuals who are either healthy or acutely ill". These findings lead those authors to suggest that patients may be wiser to appoint proxies rather than write advance directives.[10]

---

[10] Tonelli (1996) argues that due to the personhood problem, and the severe limitations of the applicability to a specific medical context of what amounts to a less than fully informed prior directive, every instructional directive would have to be validated: i.e., compared with other possible indicators of the patient's current preferences. He notes that while they are practically useful, simply "because they answer the important question of who will decide for the incompetent patient", "[t]he practical inability of surrogates to predict accurately the preferences of those they represent, however, limits the claim that these documents are effective instruments for extending individual autonomy" (p. 821). Unfortunately Tonelli did not lend convincing arguments to support his own prescription: that the instructional directive should be replaced by the "development of professional standards of medical treatment, which reflect an ethical and scientific consensus" (p. 821).

Since a competent patient *delegating* decisions would presumably expect those decisions to reflect the patient's values and preferences at the time when care decisions are made, we question whether use of any set of values, preferences or underlying preferences other than the patient's at the time of intervention would strictly meet the stipulation that the proxy act as the patient's agent in the patient's interests. Writers have questioned the legal validity and ethical justifiability of a "Ulysses contract": the prospective binding of a person's fundamental personal rights. The recognition of an incompetent person's liberty interest in refusal of treatment suggests that advance directives need to be weighed against current expressions of preferences by an incompetent person; both would have probative value as to what the *present* person would have decided if competent (Gallagher, 1996). This perspective would mean the incorporation of the advance directive *tool* within a pure best interests *criterion* which, while in most cases allowing persons to determine in advance who would make decisions or how the important decisions might be made, could provide some allowance for evidence of coercion, changed circumstances, and a change of views by the patient who has become incompetent.

Prospective Construction of the Patient

The widely discussed *thankyou theory* rationale for replacing current preferences with anticipated future preferences after treatment (Stone, 1975) lacks ethical or scientific justification (Annas & Densberger, 1984; Hiday, 1992; Venesy, 1994). As has been repeatedly demonstrated in many different contexts, hindsight is notoriously inaccurate (Fischoff, 1982; Tetlock, 1983); a current evaluation of a past event or decision may be more strategic than sincere, since the past cannot be changed, but the choice of discourse or expression of attitude to that past can effect what is to come. This is consistent with cognitive dissonance theory, which in its various applications suggests that it may be rational to make an inaccurate valuation of a position, or assessment of its risks, when the position results from a prior decision made under duress, cannot easily be changed, or when there are "sunk costs" (Brady, Clark, & Davis, 1995; Davis, 1993;

Gruppen, Margolin, Wisdom, & Grum, 1994). Post-treatment utilities would not incorporate the portion of treatment risk that has already been absorbed, and the patient may already have made costly adjustments to tolerate treatment effects (Ronis, 1992).

Weakness of the thankyou theory does not lie merely with the unreliability of the thank you, but with the fact that the thanking person is not the pre-intervention person. The thankyou theory is also particularly feeble in the case of psychiatric or psychological treatment which aims to alter the subjective realms of the patient (see Michels, 1973). Many of the most "effective" therapies involve changes in who we are (Keen, 1997; Mandell & Selz, 1996) and, at least with respect to psychiatric drugs, psychosurgery and shock treatments, often has "secondary effects" (characterized as primary by Breggin, 1991) impairing perception, memory, intellectual abilities, volition and emotion. Up to 10% of patients seen in geriatric clinics for mental impairment suffer drug-induced dementia, mostly from *prescribed* drugs (Arnold & Kumar, 1993). Hence, the patient's post-treatment "thankyou" is doubtful for reasons shared with other cases of repeated power and dominance that appear to create thankfulness or acceptance— e.g., some severe abuse victims and brainwashed prisoners (see Burstow, 1992). A pure best interests criterion based upon respect for the person necessitates construction of preferences for the present person, not for the person a proxy decision maker would like to create.

Implementation of the Pure Best Interests Objective

The Evaluability Problem

The *means* of ascertaining pure best interests must always be explicitly justified with respect to the *end*: protection of the patient's own interests as a unique individual with a liberty interest. The confusion found in the literature and in court decisions with respect to proxy decision-making is due, in our view, largely to a confusion of ends with means. Satisfying current underlying preferences has not gained acceptance as the goal of all proxy decision-making because they are, by definition, unknowable. The devices that have been adopted to generate proxy decisions, however, do not resolve the ethical problems. While they may give clear "results", how can they be evaluated? While proxy processes can be audited to ensure that they meet some operational standards, what are the standards supposed to achieve?

Despite the unknowability of the *actual* underlying preferences of incompetent persons, it should be possible to evaluate the likelihood that one preference estimation process, in general or in a particular case, will generate better estimates than another process. There is a myriad of ways to support such evaluations, including logic, common sense, and experiment (e.g. by testing an estimation process with competent persons). The pure best interests model would explicitly exclude non-patient interests and require an investigation into the patient's interests as a *unique* person. It does not, as does the substituted judgment rule, require blindly following directives or preferences expressed prior to the onset of incompetence, recognizing that because the patient's interests and circumstances have changed, the patient, if still competent, might have changed her or his mind.

The Role of Current Expressions of Patient Preferences

In psychiatric care, non-acceptance of expressed patient preferences are justified by abstractly constructing the "real" patient, having "underlying" preferences, by "factoring out" the effects of a mental illness. Does this not logically require a belief that the "real" patient is hidden away and that his or her actions or speech are controlled or influenced by an alien force? (Which raises the irony of forcibly treating a patient because of "irrational" complaints about being controlled by external forces!) How such a belief could be justified will not be dealt with here (Hall, 1996). However, bizarre beliefs and peculiar perceptions may well be potentially understandable metaphoric inventions to provide meaning, and are potentially understandable by others if an effort is made (Barham, 1986; Corin & Lauzon, 1992; Dawson, 1994; Kirmayer, 1994; Van Der Geest, 1995). "Lack of insight" is a common reason to discount patient perspectives; what then are we to make of indications that patients with lack of insight are the most dissatisfied with psychiatric care (Barker, Shergill, Higginson, & Orrell, 1996)? The danger, of course, is that to the extent that patient dissatisfaction is attributed to lack of insight, and hence discountable, the concept of patient satisfaction could have no pertinence to evaluations of mental health services.

While refusal of treatment among psychiatric patients is often seen as indicative of irrationality or lack of insight, the rate of treatment refusal has been shown to be extremely low, at less than 10% (Appelbaum & Schwartz, 1992; Brooks, 1987). Donovan and Blake (1992) note that it appears from a "proliferation of work" that non-compliance with medical treatment in general ranges in the area of 50%: "... it is difficult then to view such behaviour as deviant or even abnormal" (pp. 507, 510). One might interpret these figures as indicating that if there is psychiatric patient irrationality with respect to treatment choices, it is more often in accepting "indicated" treatment than in rejecting it. As pointed out forcefully in a recent editorial in the *British Journal of Psychiatry* (Thomas, Romme and Hamelijnck, 1996, p. 403), users are deeply dissatisfied with the role of psychiatrists:

> Users complain that all we have to offer is medication, when what they want is a job or decent housing. We take experiences rooted in social adversity and extract from them only those aspects of importance to us, the symptoms of mental illness. ... Many users regard psychiatry as an alien tongue in opposition to the languages that they want to describe their problems.

Nevertheless, regardless of the due process model for reviewing psychiatric treatment refusals (treatment-driven, rights-driven, or hybrid), "the vast majority of adjudicated cases result in the patient's objection being overruled ... " (Kapp, 1994, p. 230; see also Cournos, McKinnon, & Adams, 1988).

The Validity Problem

There is an obvious validity problem of any substituted decision process which purports to replace the *expressed* preferences of a patient with a contrary estimate of underlying preferences. Under the pure best interests criterion such a problem can and must be faced squarely. Validity problems cannot even be formulated, let alone addressed, with the best interests test and substituted judgment rules, because there is no consensus as to what those processes are supposed to accomplish. The explicit nature of the validity problem in application of the pure best interests criterion should make it much harder under a pure best interests model to ignore the expressed views, goals, fears, values and attitudes of patients designated as incompetent, than under the two traditional models.

Awareness of potential treatments and providing probability ranges of various outcomes may be largely the preserve of medical experts, but patient preferences, and the values they spring from, are subjective (Churchill, 1977; Wikler, 1988). Hence we would in the usual case assume that patients can best assess them. Preference creation draws upon information internal to an individual, i.e., her/his basic values, and external, i.e., the person's perceived circumstances. Therefore, in *estimating* patient preferences the proxy must operate on the basis of assumptions or knowledge about the patient's preferences, values and circumstances, aiming to gain familiarity with the patient as a *unique individual*. Such knowledge could be obtained by:

- having known the patient;
- researching the patient's past actions and pronouncements;
- taking into account the patient's current expressed values and preferences, since even a patient judged incompetent by a strict standard may be fairly capable to communicate how decision alternatives impact on that patient's values, likes, and dislikes; or
- interviewing persons who have known the patient (e.g., friends, relatives, caregivers), to obtain background on the patient as well as their opinions on what the patient would want.

Decisions ostensibly substituted on behalf of the patient should not be confused with interventions which represent other interests[11]. We should not expect a proxy decision maker to impose on the incompetent patient constraints or interests that the temporarily competent patient wouldn't self impose (Brock, 1996; Lynn, 1988). Advance directives, medical advice, family opinion and community values would play a role as more or less credible evidence as to what the patient's pure best interests are, to the extent that current preferences are not directly observable, but not as *substitutes* for the patient's wishes. Prior expressions of patient preferences, to the extent that they were consistent, competent, applicable to the problem at hand, and uncoerced, perhaps require extra weight: Gutheil and Appelbaum (1983, p. 9) have warned that discounting prior preferences on the basis of changed circumstances raises the risk of entry of "largely arbitrary" best interests considerations in what is ostensibly a substituted judgment.

Proxy Failure in Estimating Patient Preferences

The literature has widely acknowledged that substituted decision-making has, to date, dramatically failed with respect to the criterion of estimating current patient preferences. Emanuel E. J. and Emanuel L. L. (1992, p. 2071) were among the first to sound the alarm in response to the early empirical research:

> The recent trend has been to rely on proxy decision making for incompetent patients. Support for proxy decision making has failed to acknowledge the growing body of ethical and empirical research suggesting that proxy decision making fails to realize its objective of promoting the patient's medical care preferences. ... [Solutions are needed in the near future due to] the persistent failing of proxy decision making as currently justified and practiced.

---

[11] E.g., those of family, public safety, treatment system. See also Sinclair's [1993] comparison of organ

Lynn (1992, p. 2083) stated that "the current empirical evidence is inadequate to be of much help in guiding policy development about decision-making for incompetent adults". The sparse literature which we do have attempting to assess the accuracy of proxy and clinician predictions of patient preferences, none of which addresses psychiatric patients, is discouraging; results are "moderate" (Sulmasy, Haller, & Terry, 1994) or "poor" (Druley et al., 1993). Lynn and Teno's (1993) literature review suggests that surrogates "err substantially". Their call for further empirical research has met with little response.[12]

The Inevitability of Uncertainty

It is important to note that a pure best interests model must inevitably magnify the uncertainty implicit in achieving a substituted decision. This should be viewed as a virtue, encouraging more careful enquiry into the patient's circumstances, interests, values and background, and suggesting caution before replacing the patient's expressed preferences with a contrary estimate of "underlying" preferences (McCubbin, 1997). Criticism of a model because its results may be more uncertain than alternatives (e.g., the Smith and Nunn [1995] critique of Baergen's model) should not be too easily accepted. In addition to the moral justification for a

---

donation by incompetent persons under Jewish, best interests and substituted judgment law.

[12] Tonelli (1996, p. 818) describes a study in which even when surrogates were designated by the patient and were confident about their ability to decide what the patient would want, they "were unable to choose consistently treatment options conforming to the preferences of the patient ... substituted judgment appears to be an illusory ideal". A study of prospective decisions by proxies on behalf of their spouses with Alzheimer's Disease was revealing: faced with hypothetical care decisions in the case of critical illness or coma, while 70% of the proxies felt sure of their spouses' preferences, the decision-making criteria that the proxies chose suggested that they "anticipate making the most reasonable decision for all involved rather than basing their decision solely on a 'substituted judgment' standard" (Mezey, Kluger, Maislin, & Mittelman, 1996, p. 149). An examination by Warren et al. (1986) of proxy consent to research in a nursing home found that of the 55 proxies who believed that the patient, if he or she were competent, would have refused to consent to participate in the study, 17 of the proxies gave consent anyway.

substituted judgment model, the most important technical criterion for such a model is its *validity*: that it really measures (albeit with imperfect accuracy and reliability) that which it ostensibly aims to measure.

As noted by Baergen (1995b, p. 373), a model which seems to provide a fairly consistent high degree of certainty "has, at least, oversimplified the situation, or has misrepresented it altogether". One could, for example, blindly follow to the letter all advance directives, and be "certain". But of what? Or, if substituted decisions were to be made according to a professional judgment criterion, the "correct" decisions would more easily present themselves: but their justification would not lie in the value of individual autonomy. Again, as Baergen aptly noted: "... [S]urrogates who are highly confident about the decisions they make in complex or difficult situations are probably telling us more about themselves than about their evidence".

We have noted an unfortunate tendency in the literature whereby authors seek to ground a substituted decision-making model in paternalism, when faced with frustration in implementing autonomy values, instead of going back to the basics of clarifying and implementing an autonomy based model. Such changes in direction may solve problems for family members and clinicians, but it remains to be demonstrated that the easy paternalistic solutions are consistent with what we hold as fundamentally moral, and that such solutions are what individuals would want in the case of their eventual incompetency.

Monitoring and Evaluation

Any decision to impose interventions, including psychiatric treatment, against the expressed will of a patient deemed incompetent should be explicitly justified in a legal or quasi-legal venue (see: Carney & Tait, 1997; Coker & Johns, 1994; Rabinowitz, 1994; Sundram, 1988) such that it can be seen that a fair and appropriate process of determining the patient's preferences was carried out, that there is reasonable certainty in the resulting estimates, and that the proposed intervention can be expected to satisfy those estimated preferences (Engelhardt & McCullough, 1981). Entrusting such decisions to "professional judgment" (as allowed by the Charters standard, prevailing in the U.S.) or to unmonitored guardians does not adequately protect against bias or error, given the importance of such decisions (particularly when they are made against the expressed will of the patient) [Bopp & Coleson, 1996; McCubbin, 1997; see Buchanan & Brock, 1989].

This implies that such decisions should not be fully entrusted to individual psychiatrists, guardians, family members or other caregivers. Even if such persons could be assumed to be only concerned for the patients, proxy decisions need to juggle a complex set of factors, thereby requiring the participation of more than one individual in the process. Without supervision, there may be reason to fear that proxies will tend to use more short-cuts in the decision-making process than is justifiable, given that the incentives to make a good decision are not the same for the proxy as they would be for a competent patient[13]. Achieving effective mechanisms to monitor substituted decision-making will require careful thought however. While previous literature suggested that "situational involvement", i.e., the expectation of having to justify the decision, increases effort and the use of systematic information processing strategies, Chinburapa *et al.* (1993) did not find such results in their own study — possibly due to unexamined contingency

---

[13] See, for example, a study by Harth and Thong (1995) showing serious misunderstanding by parents of the implications of their consent to the participation by their children in clinical drug trials. See also the discussion of trust by subjects in research projects by Kass, Sugarman, Faden and Schoch-Spana (1996), and a review of recent psychiatric ethics literature by Adshead (1996).

variables, or due to the hypothetical nature of such experiments which, according to Janis and Mann (1977, pp. 69, 98), may not induce sufficient stress to simulate real conditions.

This non-finding is consistent, however, with the often disappointing results of efforts to improve decisions and accountability (e.g., Connors et al., 1995; Elliott, 1996; Jewkes & Murcott, 1996). Care has to be taken in the design of such processes to avoid distortion of objectives through cooptation and systemic intransigence (McCubbin, 1994; McCubbin & Cohen, 1997; Seidman, 1978; White, 1993, 1996; White & Mercier, 1991). The design must ensure that a new structure and process of substituted decision-making is evaluable with respect to its objectives and incorporates an evaluation program.

It should be noted that monitoring with respect to individual cases raises a different set of evaluation issues than does the testing and evaluation of a *process*. Despite daunting methodological problems, it might be possible to test in controlled or experimental conditions the relative merits of alternative preference estimation processes (see Rosenfeld, White, & Passik, 1997). Monitoring would aim not to "prove" that preferences were correctly estimated, but that procedures were followed, that there are no serious conflicts of interest, etc. The purpose of testing, developing and evaluating proxy decision-making processes, and implementing them in individual cases, would be to maximize the degree to which such processes can be expected to satisfy the pure best interests criterion.

Many jurisdictions have established mental health review boards to approve cases of psychiatric treatment or commitment made against a patient's expressed will. These boards should be developing the specialized expertise required to ensure more accurate estimates of patient preferences, matching them up with medical and other knowledge. Today, however, such boards generally operate as rubber stamps for the "medically indicated" professional judgment of an institutional psychiatrist (Vestergaard, 1994). This is not surprising, due to their makeup, their lack of due process, and the lack of scrutiny of their operations by users, public, government or researchers. In the province of Québec, for example, commitment review boards are composed of two medical practitioners and one lawyer, hearings are held only when requested by the patient, and only 25% of patients have the help of an advocate or lawyer at hearings whose interventions, if any, are not referred to in the extremely brief hearing reports (Cohen D. et al., 1997).

According to Storch (1992, p. 1034): "Physicians, not lawyers and judges, should make medical decisions". The problem, however, isn't in the location of expertise but in the location of power. It is right for medical practitioners to propose treatment, but not to decide treatment (Miller F. G., 1993; see Veatch [1995b] regarding the normative basis of clinical practice and medical decisions). There is no justification for a double standard that accepts the role of the competent patient as client who is expected to choose among alternatives according to personal values, while relegating the incompetent person to the role of passive recipient of medical paternalism.

Limits become particularly crucial where the treating psychiatrist also played a role — which Kopelman (1990) would consider to be inevitably value-laden — in the determination of the patient's incompetence or "dangerousness", as is well explained by Meisel (1979, p. 476):

> The physician could be empowered to make the decision as to whether and how the patient should be treated whenever an exception to informed consent is invoked. Since it is the physician, however, who determines whether or not conditions are appropriate for invoking an exception in the first place, such a solution would create an incentive for the physician to invoke an exception in order to provide treatment to a patient whom he believes would refuse it. Since the power to invoke an exception is the power to alter the balance between individualism and health, it is preferable that the right to make the ultimate treatment decision ordinarily not be vested in the physician since to do so would be to further tip the balance away from that established by the doctrine of informed consent — a balance favoring individualism.

Therefore, determination of competency should be independent of the patient's attitude to a medical recommendation, and should be based on functional criteria of decision-making capacities (Weisstub, 1990). Dominance on treatment review boards should rather be held by those whose training or inclinations facilitate the assessment of the client's pure best interests, independent of medical model ideology (i.e., users and ex-patients, lawyers, social workers, psychologists, ethicists and patient advocates).

Enhancing Autonomy and Competence

It has now become accepted, although less so in practice, that competency need not be defined as all or nothing (Silberfeld, 1990; Spar, Hankin, & Stodden, 1995; Winick, 1991), since individuals are seen as being multi-dimensional with diverse degrees of capacity in each dimension (Annas and Densberger, 1984; Fulbrook, 1994; Tomossy & Weisstub, 1997), which may fluctuate with time (Appelbaum & Roth, 1981) and context (Levesque, 1996). Accordingly, some jurisdictions impose proxy power in only specific areas of incapacity, and a few seek ways to "supplement" competency. As prosthetics and wheelchairs can enable physically challenged persons, various means could be found to render competent a partially incompetent patient, or to empower patients having little power (see Campbell, 1994). The most obvious example of this is putting at the patient's disposal agents, advocates or assistants who aid the patient in interpreting information regarding decision options, and help the patient develop and express preferences (see Olley & Ogloff, 1995). It is crucial to ensure that the incentives of these persons are of a nature that creates confidence that their first priority is the wishes or pure best interests of the patient (see Boudreau, 1991; Pellegrino, 1993; Simmons, 1990, pp. 232-235); there is no point in an advocate duplicating the role of a clinician, whose role is the psychiatric "best interests" of the patient (Eth, Levine, & Lyon-Levine, 1984; see Wettstein & Roth, 1988; Wolf, 1994).

Conclusion: Protection of Patient Pure Best Interests

The implication of the pure best interests criterion is that imposed decision-making must find ways to approach as closely as possible the patient's own values, and seek justification on that basis. If this cannot be done, we may have to admit that imposition of treatment against the will of a patient is unethical, unless justification lies in grounds other than the pure best interests of the patient. Responsible use of power implies bringing such grounds into the light so that they can be debated and evaluated. One merit of a *pure* best interests approach, as opposed to "best interests", is that it requires the clear separation of what the patient would want, according to that patient's basic values, from what others would want according to their values.

Aside from efficiency arguments that establish health as a public good (Evans & Stoddart, 1990; Laurell & Arellano, 1996) and ill health as an individually unpredictable eventuality that members of society might want to be insured against (see Rawls, 1971), society has an ethical obligation to go to great pains to ensure that where it imposes its power upon individuals it is done responsibly. In order to give a reasonable assurance that a substituted decision is made only with regard to the patient's interest, considerable effort has to be expended in ascertaining or constructing the patient's preferences. Obviously there are limits to the resources that should be employed in doing so, but it is clearly not sufficient to rely upon a medical model or professional orientation that abstracts the individual patient, in effect ignoring that patient as an individual in a social context, and treating that person as one of a large homogeneous group of defective biological machines (see Hewa & Hetherington, 1995; Kopelman, 1996; Silvers, 1996). In any field, "actions must sometimes be taken with far less than perfect knowledge of the resulting consequences. However, to make no estimate at all substitutes prejudice and happenstance for rationality" (Candee & Puka, 1984). Therefore, in allocating resources aimed at protecting the interests of autonomy-impaired persons we are ethically obliged to err on the side of generosity.

Actually putting into practice a complex but adequate substituted decision-making process would have to be facilitated by laws, guidelines, and judicial and administrative procedures, as well as by well-trained independent proxies. The question of feasibility is crucial to the ethical justifiability of substituted decision-making. At least in the case of patients able to express wishes, if proxy decision-making cannot be done well, it should not be done at all. Further research is needed to develop proxy decision-making procedures that incorporate a sense of their own fallibility, and which, as this article argues, can be evaluated with respect to the pure best interests objective.

This article is an outgrowth of a project which aims to assess the implications of uncertainty and bias for how proxy decision-making should be done (McCubbin, 1997). It became clear that even defining, conceptually, "uncertainty" and "bias" was impossible when the purpose of substitute decision-making was unclear. The present article, therefore, hopes to have responded to this problem by defining decisions based on the pure best interests of a person as those which can be justifiably expected to most closely resemble the decisions the patient would make if temporarily competent. This objective is based purely on the patient's ascribed autonomy values. Any method of substituted decision-making then becomes, in effect, a preference estimation process. Adoption of any such process should, therefore, be justified on scientific grounds as able to validly, reliably, and without bias estimate patient preferences.

References

Adshead, G. (1996). Ethics in psychiatric research and practice. Current Opinion in Psychiatry, 9, 348-351.

Annas, G. J., & Densberger, J. E. (1984). Competence to refuse medical treatment: Autonomy vs. paternalism. Toledo Law Review, 15, 561-596.

Appelbaum, P. S. (1992). Legal aspects of clinical care for severely mentally ill, homeless persons. Bulletin of the American Academy of Psychiatry and Law, 20, 455-473.

Appelbaum, P. S., & Grisso, T. (1988). Assessing patients' capacities to consent to treatment. New England Journal of Medicine, 319, 1635-1638.

Appelbaum, P. S., & Grisso, T. (1995). The MacArthur Treatment Competence Study. I: mental illness and competence to consent to treatment. Law and Human Behavior, 19, 105-126.

Appelbaum, P. S., & Roth, L. H. (1981). Clinical issues in the assessment of competency. American Journal of Psychiatry, 138, 1462-1467.

Appelbaum, P. S., & Schwartz, W. F. (1992). Minimizing the social cost of choosing treatment for the involuntarily hospitalized mentally-ill patient: A new approach to defining the patient's role. Connecticut Law Review, 24, 433-485.

Areen, J. (1987). The legal status of consent obtained from families of adult patients to withold or withdraw treatment. JAMA, 258, 229-235.

Arnold, S. E., & Kumar, A. (1993). Reversible dementias. Medical Clinics of North America, 77, 215-230.

Arras, J. D. (1988). The severely demented, minimally functional patient: An ethical analysis. Journal of the American Geriatrics Society, 36, 938-944.

Arrow, K. J. (1982). Risk perception in psychology and economics. Economic Inquiry, 20, 1-9.

Atkinson, J. M. (1996). The community of strangers: Supervision and the new right. Health and Social Care in the Community, 4, 122-125.

Auerbach, V. S., & Banja, J. D. (1996). Assessing client competence to participate in rehabilitation decision making. NeuroRehabilitation, 6, 123-132.

Baergen, R. (1995a). Revising the substituted judgment standard. Journal of Clinical Ethics, 6, 30-38.

Baergen, R. (1995b). Surrogates and uncertainty. Journal of Clinical Ethics, 6, 372-377.

Barham, P. (1986). Schizophrenia and human value. New York: Blackwell.

Barker, D. A., Shergill, S. S., Higginson, I., & Orrell M. W. (1996). Patients' views towards care received from psychiatrists. British Journal of Psychiatry, 168, 641-646.

Beahrs, J. O., & Rogers, J. L. (1993). Appropriate short-term risk in psychiatry and the law. Bulletin of the American Academy of Psychiatry and Law, 21, 53-67.

Beauchamp, T. L., & Childress, J. F. (1989). Principles of biomedical ethics (3rd. ed.). Oxford: Oxford University Press.

Bernstein, R. J. (1995). Rethinking responsibility. Hastings Center Report, 25(7), 13-20.

Bloche, M. G., & Cournos, F. (1990). Mental health policy for the 1990s: Tinkering in the interstices. Journal of Health and Social Behavior, 15, 387-411.

Bopp, J. Jr., & Coleson, R.E. (1996). A critique of family members as proxy decisionmakers without legal limits. Issues in Law & Medicine, 12, 133-165.

Boudreau, F. (1991). Stakeholders as partners: The challenges of partnership in Quebec mental health policy. Canadian Journal of Community Mental Health, 10, 7-28.

Boyd, N. F., Sutherland, H. J., Heasman, K. Z., Tritchler, D. L., & Cummings, B. J. (1990). Whose utilities for decision analysis? Medical Decision Making, 10, 58-67.

Brady, G. L., Clark, J. R., & Davis, W. L. (1995). The political economy of dissonance. Public Choice, 82, 37.

Breggin, P. R. (1991). Toxic psychiatry. New York: St. Martin's Press.

Brock, D. W. (1994). Good decision-making for incompetent patients. Hastings Center Report, 24(6), S8-S11.

Brock, D. W. (1996). What is the moral authority of family members to act as surrogates for incompetent patients? Milbank Quarterly, 74, 599-618.

Brooks, A. D. (1987). The right to refuse antipsychotic medications: Law and policy. Rutgers Law Review, 39, 339-376.

Brown, P. (1984). The right to refuse treatment and the movement for mental health reform. Journal of Health Politics, Policy and Law, 9, 291-313.

Buchanan, A., & Brock, D. W. (1986). Deciding for others. Milbank Quarterly, 64(Supp. 2), 17-94.

Buchanan, A., & Brock, D. W. (1989). Deciding for others: The ethics of surrogate decision making. Cambridge: Cambridge University Press.

Burstow, B. (1992). Radical feminist therapy: Working in the context of violence. Newbury Park, CA: Sage.

Bursztajn, H. J., Chanowitz, B., Gutheil, T. G., & Hamm, R. M. (1992). Micro-effects of language on risk perception in drug prescribing behavior. Bulletin of the American Academy of Psychiatry and Law, 20, 59-66.

Callahan, J. C. (1984). Liberty, beneficence, and involuntary confinement. Journal of Medicine and Philosophy, 9, 261-293.

Campbell, B. (1994). What makes empowerment so difficult? Taken Sept. 1996 from World Wide Web (http://www.SJUVM.stjohns.edu:70/00/organize/support/madness/campbell.barriers).

Candee, D., & Puka, B. (1984). An analytic approach to resolving problems in medical ethics. Journal of Medical Ethics, 2, 61-70.

Cantor, N. L. (1996). Discarding substituted judgement and best interests: Toward a constructive preference standard for dying, previously competent patients without advance instructions. Rutgers Law Review, 48, 1193-1272.

Carney, T., & Tait, D. (1997). Adult guardianship: Narrative readings in the 'shadow' of the law? Presented to the XXIInd International Congress on Law and Mental Health, Montréal.

Chamberlin, J. (1990). The ex-patient's movement: Where we've been and where we're going. Journal of Mind and Behavior, 11, 323-336.

Corin, E., & Lauzon, G. (1992) Positive withdrawal and the quest for meaning: The rconstruction of experience among schizophrenics. Psychiatry, 55, 266-281.

Chinburapa, V., Larson, L. N., Brucks, M., Draugalis, J., Bootman, J. L., & Puto, C. P. (1993). Physician prescribing decisions: The effects of situational involvement and task complexity of information acquisition and decision making. Social Science and Medicine, 36, 1473-1482.

Churchill, L. R. (1977). Tacit components of medical ethics: Making decisions in the clinic. Journal of Medical Ethics, 3, 129-132.

Cichon, D. E. (1989). The Eighth Circuit and professional judgment: Retrenchment of the constitutional right to refuse antipsychotic medication. Creighton Law Review, 22, 889-953.

Cohen, A., & Eastman, N. (1996). A survey of the use of supervision registers in South Thames (West) Region. Journal of Forensic Psychiatry, 7, 653-661.

Cohen, D., & McCubbin, M. (1990). The political economy of tardive dyskinesia: Asymmetries in power and responsibility. Journal of Mind and Behavior, 11, 465-488.

Cohen, D., & Remler, M. (1992). Should the right of mental patients to refuse treatment with psychotropic drugs be severely curtailed? A debate. In E. Gambrill & R. Pruger (Eds.), Controversial issues in social work (pp. 301-315). Boston: Allyn & Bacon.

Cohen, D., Thomas, G., Dallaire, B., McCubbin, M., Morin, P., & Fortier, R. (1997). Savoir, pouvoir, et dangérosité civile: Une étude quantitative et qualitative des décisions de révision de cure fermée de la Commission des affaires sociales du Québec,, 1975-1993. Unpublished manuscript.

Cohen, L. M., Germain, M., Woods, A., Gilman, E. D., & McCue, J. D. (1993). Patient attitudes and psychological considerations in dialysis discontinuation. Psychosomatics, 34, 395-401.

Coker, L. H., & Johns, A. F. (1994, Dec.). Guardianship for elders: Process and issues. Journal of Gerontological Nursing, 20, 25-32.

Connors, A. F., Dawson, N. V., Desbiens, N. A., Fulkerson, W. J. Jr, Goldman, L. et al. (1995). A controlled trial to improve care for seriously ill hospitalized patients: The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT). JAMA, 274, 1591-1598.

Cournos, F., McKinnon, K., & Adams, C. (1988). A comparison of clinical and judicial procedures for reviewing requests for involuntary medication in New York. Hospital and Community Psychiatry, 39, 851-855.

Davis, W. L. (1993). Economists' uses for cognitive dissonance: An interdisciplinary note. Psychological Reports, 73, 1179-1183.

Dawson, P. J. (1994). Philosophy, biology and mental disorder. Journal of Advanced Nursing, 20, 587-596.

Degrazia, D. (1995). Value theory and the best interests standard. Bioethics, 9, 50-61.

Delquié, P. (1993). Inconsistent trade-offs between attributes: New evidence in preference assessment biases. Management Science, 39, 1382-1395.

Diamond, R. J. (1995, Summer). Coercion in the community: Issues for mature treatment systems. New Directions for Mental Health Services, 66, 3-18.

Donovan, J. L., & Blake, D. R. (1992). Patient non-compliance: Deviance or reasoned decision-making? Social Science and Medicine, 34, 507-513.

Dresser, R. (1996). Still troubled: In re Martin. Hastings Center Report, 26(4), 21-22.

Dresser, R., & Whitehouse, P. J. (1994). The incompetent patient on the slippery slope. Hastings Center Report, 24(4), 6-12.

Druley, J. A., Ditto, P. H., Moore, K. A., Danks, J. H., Townsend, A. and Smucker, W. D. (1993). Physicians' predictions of elderly outpatients' preferences for life-sustaining treatment. Journal of Family Practice, 37, 469-475.

Elliott, R. L. (1996). Mental health reform in Georgia, 1992 to 1996. Psychiatric Services, 47, 1205-1211.

Emanuel, E. J., & Emanuel, L. L. (1992). Proxy decision making for incompetent patients: an ethical and empirical analysis. JAMA, 267, 2067-2071.

Emanuel, L. L., & Emanuel, E. J. (1993). Decisions at the end of life: guided by communities of patients. Hastings Center Report, 23(5), 6-14.

Engelhardt, H. T. Jr, & McCullough, L. B. (1981). Ethics in psychiatry. In S. Arieti & H. K. H. Brodie (Eds.), American handbook of psychiatry (2nd ed.) [pp. 795-818]. New York: Basic Books.

Erlen, J. A. (1995). When the patient lacks decision-making capacity. Orthopaedic Nursing, 14(4), 51-54.

Eth, S., Levine, M. L., & Lyon-Levine, M. (1984). Ethical conflicts at the interface of advocacy and psychiatry. Hospital and Community Psychiatry, 35, 665-666.

Evans, R. G., & Stoddart, G. L. (1990). Producing health, consuming health care. Social Science and Medicine, 31, 1347-1363.

Everett, B. (1994). Something is happening: The contemporary consumer and psychiatric survivor movement in historical context. Journal of Mind and Behavior, 15, 55-70.

Fennell, P. (1992). Balancing care and control: Guardianship, community treatment orders and patient safeguards. International Journal of Law and Psychiatry, 15, 205-235.

Fennell, P. (1995, 3 March). Force feeding and the Mental Health Act, 1983. New Law Journal, 319-320.

Fischhoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases (pp. 335-351). Cambridge: Cambridge University Press.

Fisher, S. & Greenberg, R.P. (1997). From placebo to panacea: Putting psychiatric drugs to the test. New York: Wiley.

Fulbrook, P. (1994). Assessing mental competence of patients and relatives. Journal of Advanced Nursing, 20, 457-461.

Fulop, N. J. (1995). Involuntary outpatient civil commitment: What can Britain learn from the U.S. experience? A civil liberties perspective. International Journal of Law and Psychiatry, 18, 291-303.

Gaffin, D. (1996). From "bugging-out" to "chilling-out": Manipulating emotion and evoking reason in a forensic psychiatric hospital.. Culture, Medicine and Psychiatry, 20, 199-228.

Gallagher, E. M. (1996, January). Advance directives for psychiatric care: A legal analysis. Presented to the Continuing Legal Education Seminar "State of Mind", sponsored by the Washington Defender's Association, Washington.

Gigliotti, G. A., & Rubin, J. (1991). The right to refuse treatment: An application of the economic principles of decision-making under uncertainty. International Journal of Law and Psychiatry, 14, 405-416.

Gordon, R. M. (1993). Out to pasture: A case for the retirement of Canadian mental health legislation. Canadian Journal of Community Mental Health, 12, 37-55.

Greenberg, J., Pyszczynski, T., Warner, S., & Bralow, D. (1994). A prognostic utility bias in judgments of similarity between past and present instances: How available information is deemed useful for prediction. European Journal of Social Psychology, 24, 593-610.

Griffith, D. B. (1991). The best interests standard: A comparison of the state's parens patriae authority and judicial oversight in best interests determinations for children and incompetent patients. Issues in Law & Medicine, 7, 283-338.

Grisso, T., & Appelbaum, P. S. (1995). The MacArthur Treatment Competence Study. III: abilities of patients to consent to psychiatric and medical treatments. Law and Human Behavior, 19, 149-174.

Gruppen, L. D., Margolin, J., Wisdom, K., & Grum, C. M. (1994). Outcome bias and cognitive dissonance in evaluating treatment decisions. Academic Medicine, 69, S57-S59.

Gutheil, T. G., & Appelbaum, P. S. (1980). Substituted judgement and the physician's ethical dilemma: With special reference to the problem of the psychiatric patient. Journal of Clinical Psychiatry, 41, 303-305.

Gutheil, T. G., & Appelbaum, P. S. (1983). Substituted judgment: Best interests in disguise. Hastings Center Report, 13(3), 8-11.

Gutheil, T. G., Bursztajn, H., Kaplan, A. N., & Brodsky, A. (1987). Participation in competency assessment and treatment decisions: The role of a psychiatrist-attorney team. Mental and Physical Disability Law Reporter, 11, 446-449.

Hall, B. A. (1996). The psychiatric model: A critical analysis of its undermining effects on nursing in chronic mental illness. Advances in Nursing Science, 18, 16-26.

Harmon, L. (1990). Falling off the vine: Legal fictions and the doctrine of substituted judgment. Yale Law Journal, 100, 1-71.

Harth, S. C., & Thong, Y. H. (1995). Parental perceptions and attitudes about informed consent in clinical research involving children. Social Science and Medicine, 41, 1647-1651.

Hermann, D. H. J. (1990). Autonomy, self determination, the right of involuntarily committed persons to refuse treatment, and the use of substituted judgment in medication decisions involving incompetent persons. International Journal of Law and Psychiatry, 13, 361-385.

Hewa, S., & Hetherington, R. W. (1995). Specialists without spirit: Limitations of the mechanistic biomedical model. Theoretical Medicine, 16, 129-139.

Hiday, V. Aldigé (1992). Coercion in civil commitment: Process, preferences, and outcome. International Journal of Law and Psychiatry, 15, 359-377.

Hill, D. (1983). The politics of schizophrenia. Lanham, MD: University Press.

Holloway, F. (1996). Community psychiatric care: From libertarianism to coercion. Moral panic and mental health policy in Britain. Health Care Analysis, 4, 235-243.

Holstein, J. A. (1993). Court-ordered insanity: Interpretative practice and involuntary commitment. New York: Aldine de Gruyter.

Ingleby, D. (1985). Mental health and social order. In S. Cohen & A. Scull (Eds.), Social control and the state (pp. 141-188). Oxford: Basil Blackwell.

Janis, I. L., & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. New York: Free Press.

Jecker, N. S. (1990). The role of intimate others in medical decision making. Gerontologist, 30, 65-71.

Jewkes, R., & Murcott, A. (1996). Meanings of community. Social Science and Medicine, 43, 555-563.

Johannesson, M. (1996). A note on the relationship between ex ante and expected willingness to pay for health care. Social Science and Medicine, 42, 305-311.

Kapp, M. B. (1994). Treatment and refusal rights in mental health: Therapeutic justice and clinical accommodation. American Journal of Orthopsychiatry, 64, 223-234.

Kass, N. E., Sugarman, J., Faden, R., & Schoch-Spana, M. (1996). Trust, the fragile foundation of contemporary biomedical research. Hastings Center Report, 26(5), 25-29.

Keen, E. (1997). Psychiatric power and levels of concealment. Journal of Constructivist Psychology, 10, 125-137.

Keshavan, M. S., & Kennedy, J. S. (Eds.). (1992). Drug-induced dysfunction in psychiatry. New York: Hemisphere.

Kline, S.A. (1992). Substitute decision-making and the substitute decision-maker. Health Law in Canada, 13, 125-128.

Kirk, S. A., & Kutchins, H. (1992). The selling of DSM: The rhetoric of science in psychiatry. New York: Aldine de Gruyter.

Kirmayer, L. J. (1994). Improvisation and authority in illness meaning. Culture, Medicine and Psychiatry, 18, 183-214.

Kong, A., Barnett, G. O., Mosteller, F., & Youtz, C. (1986). How medical professionals evaluate expressions of probability. New England Journal of Medicine, 315, 740-744.

Kopelman, L. M. (1990). On the evaluative nature of competency and capacity judgments. International Journal of Law and Psychiatry, 13, 309-329.

Kopelman, L. M. (1996). Ethical assumptions and ambiguities in The Americans With Disabilities Act. Journal of Medicine and Philosophy, 21, 187-208.

Laor, N. (1984). Common sense ethics and psychiatry. Psychiatry, 47, 135-150.

Laurell, A. C., & Arellano, O. L. (1996). Market commodities and poor relief: The World Bank proposal for health. International Journal of Health Services, 26, 1-18.

Lee, M., & Ganzini, L. (1994). The effect of recovery from depression on preferences for life-sustaining therapy in older patients. Journal of Gerontology: Medical Sciences, 49, M15-M21.

Levesque, R. J. R. (1996). Regulating the private relations of adults with mental disabilities: Old law, new policies, hollow hopes. Behavioral Sciences and the Law, 14, 83-106.

Lidz, C. W., & Hoge, S. K. (Eds.). (1993). Coercion in mental health care. Behavioral Sciences & the Law, 11(3).

Lynn, J. (1988). Conflicts of interest in medical decision-making. Journal of the American Geriatrics Society, 36, 945-950.

Lynn, J. (1992). Procedures for making medical decisions for incompetent adults. JAMA, 267, 2082-2084.

Lynn, J., & Teno, J. M. (1993). After the Patient Self-Determination Act: The need for empirical research on formal advance directives. Hastings Center Report, 23(1): 20-24.

Malloy, T. R., Wigton, R. S., Meeske, J., & Tape, T. G. (1992). The influence of treatment descriptions on advance medical directive decisions. Journal of the American Geriatric Society, 40, 1255-1260.

Mandell, A. J., & Selz, K. A. (1996). Gooey and the rat man: Dynamical stereotypies in brain and behavior. Asheville, NC: Cielo Institute.

Mazur, D. J., & Merz, J. F. (1993). How the manner of presentation of data influences older patients in determining their treatment preferences. Journal of the American Geriatric Society, 41, 223-228.

McCubbin, M. (1994). Deinstitutionalization: The illusion of disillusion. Journal of Mind and Behavior, 15, 35-53.

McCubbin, M. (1997, June). Substituted judgment: Problems in estimating the "pure best interests" of psychiatric patients considered incompetent. Presented to the XXIInd International Congress on Law and Mental Health, Montréal.

McCubbin, M., & Cohen, D. (1996). Extremely unbalanced: Interest divergence and power disparities between clients and psychiatry. International Journal of Law and Psychiatry, 19: 1-25.

McCubbin, M., & Cohen, D. (1997). Mental health policy from a systemic perspective: Orientations for strategic reform. [unpublished]

McNeil, B. J., Pauker, S. G., Sox, H. C. Jr, & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. New England Journal of Medicine, 306, 1259-1262.

Meisel, A. (1979). The "exceptions" to the informed consent doctrine: Striking a balance between competing values in medical decisionmaking. Wisconsin Law Review, 127, 413-488.

Mezey, M., Kluger, M., Maislin, G., & Mittelman, M. (1996). Life-sustaining treatment decisions by spouses of patients with Alzheimer's Disease. Journal of the American Geriatrics Society, 44, 144-150.

Michels, R. (1973). Ethical issues of psychological and psychotherapeutic means of behavior control. Hastings Center Report, 3(4), 11-13.

Miedema, B. (1994). Control or treatment? Experiences of people who have been psychiatrically hospitalized in New Brunswick. Canadian Journal of Community Mental Health, 13, 111-122.

Miller, F. G. (1993). The concept of medically indicated treatment. Journal of Medicine and Philosophy, 18, 91-98.

Miller, R. D. (1990). Hidden agendas at the law-psychiatry interface. Journal of Psychiatry & Law, 18, 35-58.

Miller, R. D. (1994). The US Supreme Court looks at voluntariness and consent. International Journal of Law and Psychiatry, 17, 239-252.

Monahan, J., Hoge, S. K., Lidz, C., Roth, L. H., Bennett, N., Gardner, W., & Mulvey, E. (1995). Coercion and commitment: Understanding involuntary mental hospital admission. International Journal of Law and Psychiatry, 18, 249-263.

Moorhouse, A., & Weisstub, D. N. (1996). Advance directives for research: Ethical problems and responses. International Journal of Law and Psychiatry, 19, 107-141.

Moreno, J. D. (1993). Who's to choose? Surrogate decisionmaking in New York State. Hastings Center Report, 23(1), 5-11.

Moskowitz, E. H. (1996). Moral consensus in public ethics. Patient autonomy and family decisionmaking in the work of one state bioethics commission. Journal of Medicine and Philosophy, 21, 149-168.

Olley, M. C., & Ogloff, J. R. P. (1995). Patients' rights advocacy: Implications for program design and implementation. Journal of Mental Health Administration, 22, 368-376.

Parry, J. (1987). A unified theory of substitute consent: Incompetent patients' right to individualized health care decision-making. Mental and Physical Disability Law Reporter, 11, 378-385, 441.

Payton, S. (1992). The concept of the person in the parens patriae jurisdiction over previously competent persons. Journal of Medicine and Philosophy, 17, 605-645.

Pellegrino, E. D. (1993). Societal duty and moral complicity: The physician's dilemma of divided loyalty. International Journal of Law and Psychiatry, 16, 371-391.

Perlin, M. L. (1990). Are courts competent to decide competency questions?: Stripping the facade from United States v. Charters. Kansas Law Review, 38, 957-1001.

Perlin, M. L. (1991). Morality and pretextuality, psychiatry and law: of "ordinary common sense," heuristic reasoning, and cognitive dissonance. Bulletin of the American Academy of Psychiatry and Law, 19, 131-150.

Perlin, M. L. (1994). Law and Mental Disability. Charlottesville, VA: Michie.

Rabinowitz, J. (1994). Determining the competency of the neediest. Journal of Mind and Behavior, 15, 157-175.

Ramsey, P. (1978). The Saikewicz precedent: What's good for an incompetent patient? Hastings Center Report, 8(6), 36-42.

Rawls, J. (1971). A theory of justice. Cambridge: Belknap Press.

Redelmeier, D. A., Rozin, P., & Kahneman, D. (1993). Understanding patients' decisions: Cognitive and emotional perspectives. JAMA, 270, 72-76.

Reed, S. C., & Lewis, D. A. (1990). The negotiation of voluntary admission in Chicago's state mental hospitals. Journal of Psychiatry & Law, 18, 137-163.

Robertson, J. A. (1985). The geography of competency. Social Research, 52, 555-579.

Ronis, D. L. (1992). Conditional health threats: Health beliefs, decisions, and behaviors among adults. Health Psychology, 11, 127-134.

Rosenfeld, B. D., & Turkheimer, E. N. (1995a). Modeling psychiatric patients' treatment decision making. Law and Human Behavior, 19, 389-405.

Rosenfeld, B. D., & Turkheimer, E. N. (1995b). Multidimensional representation of decision-making in chronic schizophrenics. Multivariate Behavioral Research, 30, 199-211.

Rosenfeld, B. D., Turkheimer, E., & Gardner, W. (1992). Decision making in a schizophrenic population. Law and Human Behavior, 16, 651-662.

Rosenfeld, B. D., White, M., & Passik, S. D. (1997). Making treatment decisions-making with HIV infection: A pilot study of patient preferences. Medical Decision Making, 17, 307-314.

Rosenson, M. K. (1993). Social work and the right of psychiatric patients to refuse medication: A family advocate's response. Social Work, 38, 107-112.

Ross, C. A., & Pam, A. (Eds.). (1995). Pseudoscience in biological psychiatry: Blaming the body. Toronto: John Wiley & Sons.

Sehgal, A., Galbraith, A., Chesney, M., Schoenfeld, P., Charles, G., & Lo, B. (1992). How strictly do dialysis patients want their advance directives followed? JAMA, 267, 59-63.

Seidman, E. (1978). Justice, values and social science: Unexamined premises. Research in Law and Sociology, 1, 175-200.

Silberfeld, M. (1990). The mentally incompetent patient: A perspective from the competency clinic. Health Law in Canada, 11(2): 33-37.

Silvers, A. (1996). (In)equality, (ab)normality, and The Americans With Disabilities Act. Journal of Medicine and Philosophy, 21, 209-224.

Simmons, H. (1990). Unbalanced: Mental health policy in Ontario, 1930-1989. Toronto: Wall & Thompson.

Sinclair, D. (1993). Kidney donations from the legally incompetent in Jewish and comparative law. Israel Law Review, 27, 588-599.

Smith, D. G., & Nunn, S. (1995). Substituted judgment: In search of a foolproof method; a response to Baergen. Journal of Clinical Ethics, 6, 184-186.

Solomon, P., Rogers, R., Draine, J., & Meyerson, A. (1995). Interaction of the criminal justice system and psychiatric professionals in which civil commitment standards are prohibitive. Bulletin of the American Academy of Psychiatry and Law, 23, 117-128.

Spar, J. E., Hankin, M., & Stodden, A. B. (1995). Assessing mental capacity and susceptibility to undue influence. Behavioral Sciences and the Law, 13, 391-403.

Stefan, S. (1992). Leaving civil rights to the "experts": From deference to abdication under the professional judgment standard. Yale Law Journal, 102, 639-717.

Stefan, S. (1993). What constitutes departure from professional judgment? Mental and Physical Disability Law Reporter, 17, 207-213.

Stefan, S. (1996). Overview of mental health law. Presented to the Conference of the National Association for Rights Protection and Advocacy.

Stone, A. A. (1975). Mental health and law: A system in transition. Washington: U.S. Government Printer's Office.

Storch, D. D. (1992). Clinical review panels versus court hearings in treatment refusal by involuntary patients. Hospital and Community Psychiatry, 43, 1033-1035.

Strudler, A. (1988). Self-determination, incompetence, and medical jurisprudence. Journal of Medicine and Philosophy, 13, 349-365.

Sulmasy, D. P., Haller, K., & Terry, P. B. (1994). More talk, less paper: Predicting the accuracy of substituted judgments. American Journal of Medicine, 96, 432-438.

Sundram, C. J. (1988). Informed consent for major medical treatment of mentally disabled people: A new approach. New England Journal of Medicine, 318, 1368-1373.

Susman, J. (1994). Resolving hospital conflicts: A study on therapeutic jurisprudence. Journal of Psychiatry & Law, 22, 107-133.

Sutherland, H. J., Lockwood, G. A., Tritchler, D. L., Sem, F., Brooks, L., & Till, J. E. (1991). Communicating probabilistic information to cancer patients: Is there ‘noise‘ on the line? Social Science and Medicine, 32, 725-731.

Tancredi, L. R., & Slaby, A. E. (1977). Ethical policy in mental health care: The goals of psychiatric intervention. New York: PRODIST/Neal Watson.

Tetlock, P. E. (1983). Accountability and complexity of thought. Journal of Personality and Social Psychology, 45, 74-83.

Thomas, P., Romme, M., & Hamelijnck, J. (1996). Psychiatry and the politics of the underclass. British Journal of Psychiatry, 169, 401-404.

Tomossy, G. F., & Weisstub, D. N. (1997). The reform of adult guardianship laws: The case of non-therapeutic experimentation. International Journal of Law and Psychiatry, 20, 113-139.

Tonelli, M. R. (1996). Pulling the plug on living wills: A critical analysis of advance directives. Chest, 110, 816-822.

Tsevat, J., Cook, E. F., Green, M. L., Matchar, D. B., Dawson, N. V., Broste, S. K., Wu, A. W., Phillips, R. S., Oye, R. K., & Goldman, L. (1995). Health values of the seriously ill. Annals of Internal Medicine, 122, 514-520.

Van Der Geest, S. (1995). Overcoming ethnocentrism: How social science and medicine should relate to one another. Social Science and Medicine, 40, 869-872.

Van Putten, T., & Marder, S. R. (1987). Behavioral toxicity of antipsychotic drugs. Journal of Clinical Psychiatry, 48(supplement), 13-19.

Vartiainen, H., Vuorio, O., Halonen, P., & Hakola, P. (1995). The patients' opinions about curative factors in involuntary treatment. Acta Psychiatrica Scandinavica, 91, 163-166.

Veatch, R. M. (1995a). Abandoning informed consent. Hastings Center Report, 25(2), 5-12.

Veatch, R. M. (Ed.). (1995b). Bioethics and contemporary philosophy of science. Journal of Medicine and Philosophy, 20(3).

Venesy, B. A. (1994). A clinician's guide to decision making capacity and ethically sound medical decisions. American Journal of Physical Medicine & Rehabilitation, 73, 219-226.

Vestergaard, J. (1994). The Danish Mental Health Act of 1989: Psychiatric discretion and the new legalism. International Journal of Law and Psychiatry, 17, 191-210.

Wang, X. T., & Johnston, V. S. (1995). Perceived social context and risk preference: A re-examination of framing effects in a life-death decision problem. Journal of Behavioral Decision Making, 8, 279-293.

Warren, J. W., Sobal, J., Tenney, J. H., Hoopes, J. M., Damron, D., Levenson, S., DeForge, B. R., & Muncie, H. L. Jr. (1986). Informed consent by proxy: An issue in research with elderly patients. New England Journal of Medicine, 315, 1124-1128.

Wear, S. (1993). Informed consent: Patient autonomy and physician beneficence within clinical medicine. Boston: Kluwer.

Weisstub, D. N. (1990). Enquiry on mental competency: Final report. Toronto: Queen's Printer.

Weisstub, D.N. (1997). Ethical and legal reflections on euthanasia. Health Law in Canada, 18, 15-22.

Wettstein, R. M., & Roth, L. H. (1988). The psychiatrist as legal guardian. American Journal of Psychiatry, 145, 600-604.

White, D. (1993). Les processus de réforme et la structuration locale des systèmes. Le cas des réformes dans le domaine de la santé mentale au Québec. Sociologie et sociétés, 25, 77-97.

White, D. (1996). A balancing act: Mental health policy-making in Quebec. International Journal of Law and Psychiatry, 19, 289-307.

White, D., & Mercier, C. (1991). Coordinating community and public-institutional mental health services: Some unintended consequences. Social Science and Medicine, 33, 729-739.

Wikler, D. (1988). Patient interests: Clinical implications of philosophical distinctions. Journal of the American Geriatrics Society, 36, 951-958.

Winick, B. J. (1991). Competency to consent to treatment: The distinction between assent and objection. Houston Law Review, 28, 15-61.

Winick, B. J. (Ed.). (1996). A critical examination of the MacArthur Treatment Competence Study: Methodological issues, legal implications, and future directions. Psychology, Public Policy, and Law, 2(1).

Wolf, S. M. (1994). Health care reform and the future of physician ethics. Hastings Center Report, 24(2), 28-41.

Error under a Pure Best Interests Model of Proxy Decision-Making

Implications for the Justifiability of Forced Treatment

Michael McCubbin

Groupe de recherche sur les aspects sociaux de la santé et de la prévention
(GRASP), Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal
(Québec), Canada
H3C 3J7

[redacted]

A presentation based upon this paper was made to the XXIInd International Congress on Law and Mental Health, Montréal, June 1997.

13 December 1997

Abstract—It is increasingly recognized in the mental health law and ethics literature that incompetence is a necessary, but not sufficient, condition for treatment against the will of a patient; treatment also has to be in the *interests* of the patient. Using decision and expected utility theory and findings, this article raises basic ethical, logical, and methodological problems in the determination of the "pure best interests" of a patient considered incompetent, in effect a process of estimating the patient's "underlying" preferences (McCubbin and Weisstub, 1998).

Pure best interests assessment has severe problems of validity, bias and reliability. In order to be ethically justifiable, substituted decision-making needs not only to *validly* estimate the patient's own values and underlying preferences and avoid systematic *bias*, but also to be *reliable* enough such that the patient's attributed expected net benefits of a decision clearly outweigh alternatives — including the alternative not to treat.

Proxy decisions must consider, along with the risks of various treatment options and estimated patient attitudes to those risks, the weakness of the preference estimates themselves, as containing some distribution of error: deviation from actual preferences. This distribution, providing a measure of the uncertainty inherent in substituted decision-making, must also be estimated.

Given these problems, limitations upon the feasibility and justifiability of substituted decision-making are discussed. The introduction into the analysis of uncertainty greatly reduces the justifiability of substituted decision-making, especially where imposed against the expressed wishes of a patient.

Acknowledgments—I appreciate the thoughtful comments on various drafts from Michael Perlin of New York Law School, and David Cohen, Marie-Claude Roberge, and Bernadette Dallaire, at GRASP. I thank Ian Greene, of the Public Policy and Administration Program at York University, for his comments on a very early version of this paper. Dr. Greene is a superb educator whose courses in statistics and evaluation, firmly grounded in conceptual principles, made this article possible. Generous financial support has been received over the last several years from the Université de Montréal, through bursaries and research contracts.

Error under a Pure Best Interests Model of Proxy Decision-Making

Implications for the Justifiability of Forced Treatment

Introduction

A growing number of bioethics writers argue that *incompetence is a necessary but not sufficient condition for the imposition of medically recommended treatment* — the professional judgment of a medical practitioner does not assure that the patient's interests will be protected (Griffith 1991; Hermann 1990; Robertson 1985; Strudler 1988). In psychiatry, this would represent a major departure from present practice, since ordinarily the purpose of declaring medical incompetence is to impose a predetermined medical treatment, as illustrated by the fact that the primary evidence used to assess medical incompetence is often, in effect, the patient's disagreement with the psychiatrist's proposed treatment (Appelbaum and Grisso 1995; Perlin 1990; Venesy 1994).

Hence in recent years there has been growing study (although dominated by non-psychiatric issues) of how medical and other decisions on behalf of incompetent patients should be arrived at. A consensus seems to have developed around hierarchy or hybrid approaches to substituted decision-making9, whereby advance directives, substituted judgment rules, or best interest criteria are applied according to the circumstances of the individual case — in particular, the degree of information available regarding the patient's "true", "underlying", or previously expressed preferences.

---

9. See: Baergen (1995a and 1995b), Brock (1994), Gordon (1993), Kline (1992), Tomossy and Weisstub (1997).

The Pure Best Interests Criterion

McCubbin and Weisstub (1998) have argued that the paradoxical presence in the substituted decision-making hierarchy of both paternalistic and autonomy values, presumably contingent upon the presence and nature of information about the patient's competent views, renders the enterprise of proxy decision-making incomprehensible in terms of the ethical values upon which the process is based. They called for a clear unifying principle, based upon one widely shared fundamental value, that of individual autonomy. At a metaphysical level this principal derives its justification from the idea of human dignity — *respect* for the person as a person. Balancing beneficence against autonomy on *empirical* grounds is putting the cart before the horse.

McCubbin and Weisstub therefore developed the "pure best interests" criterion for substituted decision-making which aims to provide, in theory at least, a more integrated, intuitively sensible and ethically transparent approach building upon the best interest and substituted judgment rules to advance to the extent possible patient self-determination and, importantly, to protect the interests of the patient from the interests of others. In this paper as in that, *pure best interests* is understood, first, as an *objective*, rather than as a rule of procedure: to make the decision that the *present* patient would have made if temporarily competent.

This means that although a patient is deemed incompetent, her/his "underlying" preferences and values provide the only justifiable basis for substituted decisions affecting that person. The concept of "underlying preferences" serves as a myth, or regulatory ideal, aimed at maximizing autonomy and minimizing dignity offending imposition where the patient has impaired autonomy (see Laor 1984).

Operationalizing the objective of pure best interests for an incompetent person by conducting an appropriate and careful inquiry into that person's underlying preferences is a way of answering Strudler's (1988) call for recognition of an incompetent person's right to self-determination; properly done, rather than making an "abstract metaphysical claim", it "creates a tool for protecting the interests of an incompetent" person (p. 357).

Decisions ostensibly substituted *on behalf* of the patient should not be confused with other interventions which represent not the patient's interests but, for example, those of community (e.g., burdens to family, danger to others, public treatment costs). We should not expect a proxy decision maker to impose on the patient constraints that the temporarily competent patient wouldn't self impose. Advance directives, medical advice, family opinion and community values would play a role not as *substitutes* for the patient's wishes, but as *evidence* — which may not always be credible — as to what the patient's pure best interests may be. The *means* of ascertaining pure best interests must always be explicitly justified with respect to the *end*: protection of the patient's own interests as a unique individual.

The pure best interests model could therefore replace the "hierarchy" or "hybrid" composed of best interest and substituted judgment rules with an objective which explicitly excludes non-patient interests and requires an investigation into the patient's interests as a *unique* person. It does not, as does the substituted judgment rule, require blindly following directives or preferences expressed prior to the onset of incompetence, recognizing that because the patient's interests and circumstances have changed, the patient, if still competent, might have changed her or his mind.

Since the pure best interests objective necessarily requires estimating what the current patient's preferences would be under current circumstances, thereby operationalizing the objective, proxy decision-making becomes potentially evaluable according to the criterion of pure best interests. The best interests and substituted judgment processes are, at best, difficult to evaluate because their objectives are ambiguous; even where those rules seem to be followed it would not be clear what was accomplished.

The Literature on Error and Uncertainty in Proxy Processes

While a small amount of research has looked at some choice implications for proxy decision-making of the risks of various treatment or protection alternatives[10], very little attention has been paid to the uncertainty involved in *estimating* preferences. Only a few generally limited studies, which do not deal with psychiatric patients, have assessed the accuracy of proxy decision-making, as observed by Sulmasy, Haller and Terry (1994):

---

[10] The literature concerned with risks of choice alternatives in proxy decision making includes Crystal and Dejowski (1987) and Gigliotti and Rubin (1991). Lee (1994) outlined a risk analysis model of involuntary civil commitment with probabilities and costs of errors in wrongly committing and wrongly failing to commit.

A similar analysis has not been applied to error in proxy decision processes, with the notable exception of the cost minimization model of Appelbaum and Schwartz (1992). This model seeks to design optimal decision processes, in terms of the patient's role in decision-making, by minimizing a cost function incorporating the expected values of decision error costs, the costs of making a decision, and costs due to loss of patient autonomy. Since the model includes not only costs borne by patients but also costs borne by others (medical staff, proxy decision makers, social costs, etc.) it is not a pure best interests model, but rather a best interests model which may incorporate patient participation in decision-making.

While a cost minimization approach might seem rather cold-blooded, at least it aims at *explicit* evaluation of the constraints and criteria behind the adoption of a decision and decision processes, and therefore conceptual separation of the interests of patients from those of others. A key implication of that model is that reduction of the patient's role entails increased uncertainty for proxy decision makers; in taking more or less of the patient's decision-making role there is a trade-off between error costs incurred by incompetent patient decisions and error costs resulting from proxy decision error.

While Appelbaum and Schwartz sought to make the degree of attributed incompetence in effect a *result* of the calculations, their analysis also implies that incompetence defined *a priori* does not necessitate the replacement of the patient's choices with those of a proxy; rather, due to the error that could be made by proxy decision makers, incompetent patients should be given some degree of control or influence in the decision.

> Little is known ... about how accurately surrogate decision makers can make substituted judgments about what the patient would actually have chosen in a given clinical situation. The limited research in this field has suggested that substituted judgments are made with only fair to poor accuracy. Little is known about what factors might be associated with improved accuracy. Most previous studies have investigated only geriatric patients (p. 432).

Those authors found surrogates in Maryland to be, at best, "moderate" predictors of patient preferences. Druley et al. (1993) also note the lack of literature identifying factors that may improve substituted judgments and found physicians to be "poor" preference predictors. Lynn and Teno (1993) cite "a flurry of articles" which document that "surrogates (next of kin, usually) err substantially in predicting patient preferences" (p. 21). However, Lynn (1992) remarks that there are no studies of the process of substituted decision-making where a patient has neither an instruction directive nor a designated proxy.

Appelbaum and Schwartz (1992) stated that "any approach to the choice of a decision-making model that offers no means for capturing the relative importance of errors is seriously incomplete" (p. 444). Nevertheless, coercive mental health systems have paid little or no attention to the implications of error in formal or *de facto* substituted decision-making, which bear not only on *who* should make substituted decisions and on *how* they should be made, but also, as this article emphasizes, on *whether* they should be made.

Evaluation under a Pure Best Interests Model of Proxy Decision-Making

The small amount of research reported so far — and that having only tenuous application to patients subjected to forced *psychiatric* treatment, perhaps the most vulnerable class of patients — suggests that in general proxy decision-making falls far short of satisfying the pure best interests criterion. Since this article views proxy decision-making under that criterion as necessarily based upon an estimation of the "underlying" preferences of the patient, a process set up under that criterion can be *evaluated* in terms of validity, bias, reliability, feasibility, and, therefore, justifiability. The McCubbin and Weisstub (1998) article addressed theoretical validity and the ethical *a priori* justifiability of the pure best interests criterion.

This article pursues the research agenda set out in that article, to develop evaluation criteria for a pure best interests model of substituted decision-making. Accordingly, this article discusses the major issues that would arise in any proxy decision making process which aims to satisfy the pure best interests criterion: issues of bias, reliability, feasibility and, given those issues, implications for the justifiability of forced treatment.

In general, this article elaborates the simple proposition that while a patient might choose the wrong alternative due to the uncertainty faced by that patient — limited information as to costs/benefits and likelihood of outcomes — the decision facing a proxy decision maker is likely to be marked by much greater uncertainty, due to additional error involved in estimating preferences. This implies limitations or needed improvements in actual proxy decision-making processes, particularly in psychiatric care where most patients are able to express what they assert as being their own desires.

This article will not resolve the fundamental validity problem under a pure best interests model of any substituted decision process which purports to replace the *expressed* preferences of a patient with a contrary estimate of "underlying" preferences. This controversial issue has been addressed, in effect, by a large body of literature justifying or critiquing the concept and determination of incompetence and the imposition of forced treatment.

Hence, this article leaves as an open question whether any pure best interests proxy process will be invalid due to false premises ostensibly justifying the application of such a process in the abstract. Adequately satisfying the pure best interests criterion, given uncertainty, is therefore a necessary but not sufficient ethical condition for the imposition of forced treatment. Imposed decision-making must, at a minimum, find ways to base itself on the patient's true values.

If this cannot be seen to be done, we should admit that imposition of treatment against the will of a patient — even one considered incompetent — is unethical, unless justification lies in grounds other than the pure best interests of the patient. Responsible use of power implies bringing such grounds into the light so that they can be debated and evaluated.

Bias Problems: At Least on Average are the Guesses Right?

Self-Confirming Hypotheses

Treatment Refusal. A major bias in determining interests arises systematically and continually because, as noted above, very often the medical incompetence of the patient is construed from the patient's refusal of a "medically indicated" treatment. Those who initially propose and finally impose psychiatric treatment are often the same practitioners whose expert opinion establishes the patient's incompetence (if indeed there is a formal determination of incompetence), resulting almost invariably in the imposition of the pre-determined treatment. If the decision-maker is such a practitioner, one therefore has to expect a severe bias in assessing what the patient would "really" want, favouring rather those treatments preferred by psychiatry — today almost always psychoactive drugs.

Deviancy. This bias is magnified to the extent that diagnosis consists of nothing more than identifying and classifying behavioural or attitudinal deviancy, as opposed to diagnosis drawing on deviancy as merely one of a number of symptoms indicating a "real" mental illness (see Borges 1995). A finding of incompetence based upon apparent irrationality linked to such deviance should suggest as problematic an assumption that the patient's "real" values are "normal", at least with respect to the assessed "underlying" values used to justify the selection of the psychiatric treatments to which the patient objects. Once cannot but view as absurd the situation this could give rise to, where the treatment decision is based on the assumption that the patient has normal values yet it is imposed because of the patient's opposition ensuing from *non*-normal values.

The Diagnosis/Compliance/Insight/Incompetence/Decision Chain. Therefore, the replacement of patient choice by psychiatric recommendations, justified by observation of behavioural abnormality, might systematically leave the patient's pure best interests far from being met. In the extreme case, where diagnosis, incompetence assessment and treatment selection rely on nothing but the correction of deviant values, this changes from a problem of bias — systematically inaccurate estimation favouring certain classes of outcome — to one of invalidity — *estimating the wrong thing.*

McEvoy, Aland Jr, Wilson, Guy and Hawkins (1981) unwittingly described a tautological diagnosis —> treatment —> compliance —> insight —> diagnosis chain where lack of insight becomes a diagnostic factor. Citing prior research they reported:

> ... [A] patient was considered to have insight if he acknowledged some awareness of having emotional illness, and to lack insight if he vigorously denied the fact that he was disturbed. *Lack of insight was so common among schizophrenic patients that it proved to be a discriminating variable in favor of making a diagnosis of schizophrenia.* .... [I]nsight ... was related to apparent drug compliance, as measured by a patient's regular attendance at scheduled medication appointments and his *lack of complaints about medication* (p. 856, emphasis added).

The authors then report their own supporting findings:

> ... [T]he majority of our chronic hospitalized schizophrenic patients approach the question of drug compliance with no realization that they have any illness or need for treatment. .... Our evidence suggests that many of our severely ill chronic schizophrenic inpatients are not competent to decide whether they need medication (p. 857).

The "test" for such competence was geared to "insight", which was assessed on the basis of what the patients said about their illness and medication[11]. It is not clear in the article how many of the patients were involuntarily committed and forced or coerced to take their medications. The act of coercion would likely augment the patients' "negative" attitudes toward their treatment.

Non-Falsifiable Diagnoses. The severest threat to the logic inherent in these findings is the attributed importance of a high rate of "lack of insight" among a population of inpatients, when many may be inpatients *because* they disagreed with the diagnosis or treatment plan. The danger of circular logic, or "self-confirming hypothesis", in psychiatry is well described by Reich (1984). To the extent that a diagnosis is not falsifiable (Popper, 1979), it can easily become what Reich describes as a "catch-22": any possible behaviour or treatment outcome confirms the diagnosis. In order to avoid a tautological justification for forced treatment, the test for incompetence must be independent of a diagnosis of mental

---

[11] Interestingly, however, *all* patients studied had been taking neuroleptics for at least three months prior to the study, and "Improvement with medication did not correlate with insight" (p. 857). One must wonder what "improvement" consisted of, and why "lack of insight" was attributed to the diagnosis rather than to the effects of imposed neuroleptic drug treatment.

illness, of the patient's values, and of the treatment decisions favoured by caregivers and the patient[12].

Sociodemographic Value Biases

Whether or not the patient's values are normal, those of persons making or influencing substituted decisions may not be. For example, psychiatrists' very high levels of income and education are among a number of important demographic and cultural distinctions between psychiatrists and others in society, differences which become more marked with respect to often poor severely ill chronic psychiatric patients[13]. There are also systematic differences in interests and outlook between those who (successfully) chose to pursue the field of medicine and others, due to initial selection, professionalization and roles[14].

Where physicians play the major role in substituted decision-making, most of the same sources of diagnostic bias can also be expected to affect preference estimates. Scott, Shiell, and King (1996) found that the general practitioners they studied tended to test more and prescribe less with patients of high compared to low socioeconomic status. Rabinowitz (1993) reported findings from the literature of diagnostic biases related to factors such as setting, collegial opinion, culture, and physical attractiveness, and showed how bias and reasoning errors can enter in various stages of the decision-making process.

A study by Tsevat et al. (1995) with a large number of seriously ill patients (1438), their surrogates and their physicians, found that patients' self-assessed health values and health ratings exceeded those estimated by surrogates and physicians. The authors therefore suggest that "simply substituting a surrogate's

---

[12] See: Appelbaum and Grisso (1995); Grisso, Appelbaum, Mulvey, and Fletcher (1995); Greenberg and Bailey (1994); Kopelman (1990); Meisel (1979); Pruchno, Smyer, Rose, Hartman-Stein, and Henderson-Laribee (1995).

[13] See: Campinha-Bacote (1994); Fabrega Jr (1992); Koffman, Fulop, Pashley, and Coleman (1997); Littlewood (1992); Loring and Powell (1988); Rodwin (1994); Rogler (1993); Wade (1993).

[14] See: Bassford (1990); Eppard and Anderson (1995); Freidson (1984); Haas and Shaffir (1987); Kaufman (1995); Martin, Arnold and Parker (1988); Stein (1990).

preferences when a patient cannot or will not participate in decision-making may be far from an ideal practice ..." (p. 518).

Conflicts of Interest

Apart from proxy biases, many sources of information about patient preferences will introduce bias into the estimate. Background, perception and views from those who have known the patient may be skewed by the specific contexts of their interaction with the patient, or adjusted in accordance with their own values or interests. What individuals say and do is highly dependent upon current contextual considerations, and may be strategically calculated to mislead in order to obtain a benefit (Cornes and Sandler 1986). While families and institutions are likely to profess that their only concern is the patient's welfare, most also have other interests in the results of the decision, or subscribe to particular treatment ideologies which are less than fully compatible with the patient's belief system[15].

The proxy must take with a rock of salt views and information coming from someone with interests apparently not convergent with those of the patient. If the proxy decision-maker were able to identify the likely existence, direction and approximate degree of biases, then they would present less of a problem as the estimation process could try to adjust for such estimated bias.

---

[15] Regarding conflicting interests between family members and patients, see: Bursztajn and Brodsky (1994); Cohen D. and McCubbin (1990); Lesemann and Chaume (1990); Meisel (1979); Parry (1987); Redding (1993). Regarding conflicts between treatment ideologies and patient belief systems, see: Cohen C. I. (1993); Mohr (1995); Scheid (1994).

Thorough research, non-interested proxy decision makers, and independent monitoring of the substituted decision-making process will increase the likelihood that biases can be estimated and that remaining unidentified biases will cancel each other out.

Where those controlling a proxy decision-making process are interested parties — family members, practitioners, and institutional employees where the patient resides — major systematic biases will likely remain, resulting in interventions favouring those interests rather than the pure best interests of the patient.

Attributions of Attitudes to Risk

In order to avoid a pure best interests process becoming in effect professional judgment or a paternalistic view of what *others* think would be good for the patient, preference estimation must incorporate the patient's attributed *subjective* attitude to risk, and focus upon the *current* patient rather than the supposed patient at another period of time. Failure to do so will reduce the accuracy of preference estimates, and perhaps also systematically bias them if those who make proxy decisions substitute implicit or explicit attitudes to risk that either reflect values differing from those of patients or, perhaps worse, reflect the fact that costs of imposed treatment tend to be borne much more by the patient than by those charged with responsibility for decisions affecting the patient (see Cohen D. and McCubbin, 1990).

Preferences for Outcomes vs. Preferences for Choices

It is crucial to note the distinction between preferences for treatment *outcomes* and preferences for treatment *choices.* The latter incorporates not only valuations of outcome alternatives, but also assessments of the probabilities of the various outcomes (risk), attitudes to risk (based on how changes of different sizes from the status quo would be valued) and time discounting preferences (the rate at which a better tomorrow compensates for sacrifices today, and vice versa[16].

Even if proxy outcome evaluations and risk probabilities are the same as those that would be assessed by the temporarily competent patient, expected utility functions incorporating different risk attitudes can yield different treatment choices (Gigliotti and Rubin, 1991). Redelmeier and Heller (1993) found that "different techniques for assessing time preferences can lead to significantly different results" (p. 216).

Controlled Risks

Furthermore, a key consideration in considering forced treatment or treatment which reduces the patient's options, is that individuals prefer risk they can control to some lesser amount of risk they cannot control (Klein and Kunda 1994; Mitchell 1990). Hence, what might appear to be a patient's "irrational" assessment of risk probabilities might in fact reflect *attitudes* to risk, a matter of valuation rather than of fact.

Hayes (1992) notes that the risk assessor may fail to consider factors of importance to the risk *taker*, suggesting therefore that the exclusive power to determine risk needs to be wrestled away from an external expert. Also, since the *probabilities* of outcomes are often heavily influenced by social and personal factors, even here the patient as client has "expert status" (Rifkin 1994): both in estimating the odds — described by Wulff (1981) as necessarily subjective —and in altering them[17].

---

[16] See: Ellis (1993); Hilden, Glasziou, and Habbema (1992); Hammitt (1993); Loewenstein and Prelec (1993); McKenzie (1991); Olsen (1993).

[17] See: Becker and Nachtigall (1994); Chewning and Sleath (1996); Furnham (1994); Lupton (1993); Pellissier, Hazen, and Chang (1996); Ronis (1992); Stambolovic' (1996).

The Impact of Coercion

Therefore it is important to realize that *the very act of coercively replacing a patient's expressed preference with a contrary "estimated" preference will alter outcomes.* Patient choice and control has a crucial influence on achieving therapeutic/rehabilitative objectives[18], and the imposition of incentives and constraints may "crowd in" or "crowd out" intrinsic motivation (Frey 1994). Substituted decision-making must be *self-conscious*, to avoid constructing a model of reality which becomes unreal as a result of the act and implications of that very construction.

Those who would make decisions *for* a patient therefore need to consider not only "objective" medical knowledge but also matters *subjective* to the patient regarding:

- the *desirability* of alternative outcomes;
- their *probabilities*, recognizing that the patient may have a role in assessing and affecting them;
- the patient's willingness to accept *risk;*
- the patient's *temporal attitudes* to alternative benefits to be enjoyed and costs to be incurred in the future; and
- the impact of *imposing* decisions upon patients.

Ignoring such subjective factors implies that it is not the interests of the patient as a unique individual that are assessed but rather of an abstract, stereotyped and contextless medical object, hence rendering the preference estimation process at best biased and unreliable and at worst completely invalid.

---

[18] It is somewhat perverse to characterize as therapeutic, interventions which remove control from patients whose problems are at least partly related to feelings of lack of power, control or status (Syx, 1995; Winick, 1997). The role of such factors in health has been discussed extensively. See, e.g.: Antonovsky (1993); Carpentier-Roy (1995); Frank R. H. (1985); McCubbin (1997); Miedema (1994); Rosenfield (1992); Sullivan (1992); Vatz and Weinberg (1990); Wolf-Branigin and Sawilowsky (1994).

The Importance of Estimate Reliability for a Pure Best Interests Model

Rarely, if ever, have proxy decision models attempted to incorporate in their construction an awareness of their fallibility in terms of the reliability of what they try to estimate and, consequently, devise procedures aimed at appropriately responding to such fallibility. This is understandable because, as argued by McCubbin and Weisstub (1998), it was not usually clear what, if anything, the traditional models were trying to estimate. Uncertainty entered the picture in terms of whether or not standard operating procedures were followed, not in terms of the degree of reliability engendered in following those procedures. Hence applications of traditional models could be *monitored* but they could not be *evaluated.*

Since a pure best interests model has a standard of the desired outcome which is conceptually independent of the means used to achieve it — the decision the patient would make if temporarily competent — it becomes possible under such a model to rigorously and explicitly identify the nature of some of the uncertainty incurred in operationalizing it, and incorporate procedures aimed at managing or appropriately acknowledging that uncertainty. Hence this part discusses the nature of uncertainty under a pure best interests model, and the next part explores how a pure best interests model might respond to such uncertainty.

Preference Variations in Populations. There is strong recent evidence of wide variation, even among "similar" patients, in their health attitudes, self-ratings of health status, and preferences for medical interventions. These variations are not predictable from sociodemographic cohort variables (Nease et al., 1995; Sorum, 1995). Furnham (1994) found that patient perceptions of health and recovery could more easily be predicted from their other attitudes and beliefs than by demographic variables.

Similarly, Tsevat et al. (1995) found that "health values of the seriously ill vary greatly from patient to patient" and that apparently their health values "cannot readily be predicted from clinical, demographic, or health status information ... This implies that the average preference of a group should not be used in a decision involving an individual patient ..." (p. 518).

Error Distributions. Therefore, in making substituted decisions not only the risks of alternative decision outcomes have to be considered, but also the possibly large degree of *uncertainty* (Brady M. E., 1993) of the estimate itself: the degree to which, after adjustment for bias, the estimated preferences may be presumed to be incorrect. To conceptualize this statistical problem, imagine possible degrees of preference for one option compared to another arrayed on a continuum from "strongly disfavour" through "ambivalent" to "strongly favour". The actual preference is at an unknown point on the continuum.

Assuming that the estimator (the estimation process) has been adjusted for bias, and that in hypothetically repeated estimations the estimate errors are randomly distributed around the true preference, we could plot a two-tailed probability density function in the form of a bell curve. From this could be read the probability that the real value of the preference falls outside a given range of estimates. If we knew the nature of the error distribution, we could take uncertainty into account by saying, for example: "Although we have estimated that the patient would "strongly" prefer A over B, there is a 30% probability that B would actually be preferred over A".

Estimate Error Variance. Therefore, even if the preference estimator is unbiased, it is possible that the preference will be wrongly estimated due to random error — uncertainty — resulting in the wrong substituted decision. Some measure of uncertainty such as the estimate error variance has to be estimated, which could be used to provide a degree of confidence that the estimate preference or resulting decision is correct. This would allow the proxy to believe that an estimate is liable to deviate by no more than x from the correct preference, y% of the time, or that there is a z% probability that the resulting decision is wrong (see Carmines and Zeller, 1986).

Problems in Estimating Error Distributions

The literature has barely discussed the problem of reliability, leave alone develop some kind of norm, rule of thumb, or contingency rules for characterizing the degree of uncertainty inherent in preference estimates — other than too often acting as if uncertainty doesn't exist or doesn't matter. However, there is a challenging methodological problem in developing reasonably accurate empirically based measures of uncertainty.

Firstly, there is of course no "gold standard" to immediately validate current preference estimates (Lynn 1992; Sorum 1995). Secondly, while one might hope to estimate the error variance by comparing historically how past preference estimates fared against subsequent observed values, the true values will never be revealed. If a patient was incapable of formulating a preference at the time of intervention, once the patient becomes cogent (if ever), he or she might either express current preferences under now changed circumstances, or an estimate of what her or his preferences would have been if the patient were temporarily competent at the time of intervention.

Hindsight. As pointed out by McCubbin and Weisstub (1998) in their discussion of the temporal construction of the patient, such current or hindsight preferences may well be biased or unreliable indicators of previous preferences as a result of the passage of time, due to the impacts of the treatment including iatrogenic effects, or due to the patient's strategic considerations affecting honest revelation of preferences.

Furthermore, an estimate of error variance based largely on study of those who become cogent after treatment might be overly weighted with "thankyous" — observations of "regret" will be missing in the cases of those who suffer severe iatrogenic illness such as tardive dyskinesia or memory loss, or for whom treatment has been ineffective, and who remain non-cogent. Hence it is very difficult to rely on experience in order to estimate the risk that the proxy's preference estimates are wrong.

Subjective Measures of Uncertainty. Nevertheless, it is undoubtedly better to use a poor measure of uncertainty than none at all; hence efforts have to be made to provide a prudent and justifiable "best guess". In this regard Bayes' rule, employing subjective probabilities, is promising[19].

However, the issue will arise as to whose subjective perspective should be employed. For example, a patient might reasonably have much less confidence in a psychiatrist's reliability in determining the patient's preferences than the psychiatrist would have. In general, people tend to underestimate the uncertainty they face, and perhaps much more so when they are not the persons who must bear the consequences of wrong decisions (Cohen and McCubbin, 1990). We have seen, for example, that physicians underestimate the uncertainty inherent in their practice (Christensen-Szalanski and Bushyhead, 1988). Gerrity et al. (1992) state that

> ... denial of uncertainty allows physicians to make potentially threatening situations more understandable and controllable, thus enabling action to take place. Denial of uncertainty, defenses against criticism and management of patients to maintain dominance ... may be predisposing factors behind overtreatment and iatrogenesis (p. 1029).

---

[19] See: Fischoff and Beyth-Marom (1988); Gustafson, Sainfort, Johnson, and Sateia (1993); Howson and Urbach (1991); Wolf, Gruppen and Billi (1988). One problem for Bayesian analysis, given the multiple sources of evidence, is that new evidence applied to the likelihood function may not be fully credible, or may conflict with prior evidence; application of fuzzy set theory might therefore be appropriate (see: Chacko, 1991, pp. 127-151; Smithson, 1987).

A more fundamental problem with respect to the use of "science" in exercises of power over others is the ease with which human elements are camouflaged or ignored by techniques loaded with bells and whistles which seem impressive but come under inadequate scrutiny because of their specialized nature. Highly mathematical models impacting on human lives must be forced to incorporate — or adapt to — forms of knowledge which are difficult or impossible to express in the same terms as those models. As noted by Gerrity, Earp, DeVellis, and Light (1992),

> ... [H]ighly rational models such as expected utility theory and Bayesian probability theory ... describe or predict but a portion of actual behavior ... because they ignore social, cultural, and organizational forces. ... The sociology of uncertainty, with its attention to norms, beliefs, rituals, and institutional responses, has yet to be written (p. 1023).

Such denial of uncertainty may be resulting in a large number of unacknowledged errors in proxy decision-making. In one study, even when surrogates were designated by the patient and were confident about their ability to decide what the patient would want, they failed to consistently choose those treatment options conforming to the patients' preferences (Tonelli, 1996).

In general, given the lack of precise and objective measures of uncertainty, it seems more pertinent at this stage in the development of proxy decision processes to ask *whose* perspectives on uncertainty will count rather than *how* measures of uncertainty should be calculated.

Interaction of Risk, Error, and Attitudes to Risk. The above discussion points out that the region of uncertainty in substituted decision-making that is due to the risks of alternative outcomes becomes greatly enlarged by the addition of the uncertainty associated with estimating the patient's pure best interests. However, the uncertainty resulting from outcome variance and preference estimation error *combined* may be further magnified, since patient attitudes to risk (also uncertainly estimated) should be applied to *both*, and the risk utility function could be non-linear (Loehman 1994).

For example, a person who buys both lottery tickets and fire insurance may be a risk lover to some degree of risk, but with more at stake be extremely risk averse. Individuals may through frequent choices accept gradual increments of risk, but not accept the same amount of risk as an "all or nothing" proposition. Further work is needed to fully integrate both types of variance (risk and estimation error) within risk utility models (see Hey and Orme 1994), in order to cope with a situation in which the dimensions of risk, attitude to risk, and uncertainty are necessarily interactive.

Prudent social scientists realize that predicting the behaviour of an *individual* is hazardous, and ascertaining motives for that behaviour even more so. Certainly no less prudence is called for when estimating the "true" preferences of an individual while dismissing those implied by that person's behaviour, a task piling uncertainty upon uncertainty.

Decision Rules for the Management Of Uncertainty

A fundamental problem in meeting the pure best interests of persons considered incompetent is a "lack of insight" by those with proxy power as to the uncertainty necessarily involved in their decisions. To the extent that the costs of a wrong decision are not borne by the decision-maker, that person has an incentive to quickly make a decision, perhaps reducing the perceived or expressed uncertainty. This is likely to reduce the reliability of proxy decision-making unless there are mechanisms or incentives in place to ensure that alternatives are more carefully considered when there is greater uncertainty or error costs.

Wolff (1989) noted that "little is currently known about [the] clinical reasoning process" (p. 106), leaving unexplained frequently observed variations in medical care utilization. The void in knowledge is much deeper with respect to involuntary psychiatric treatment; Anderson and Eppard (1995) remarked that "The actual process of clinical decision-making during assessment for involuntary psychiatric admission has not been researched or described" (p. 727). In their brief psychophenomenological study those authors describe the admission process as relying largely on factors such as intuition and experience, with apparently little in the way of guidelines to resolve uncertainty, other than the legal requirement in Ohio to use the "least restrictive alternative".

Such criteria for decision-making under uncertainty have to be further developed and operationalized, and their presuppositions made *explicit,* to enable an ethical and reasonably good decision to be made given the estimated preferences and the estimated probabilities that decision alternatives are wrong.

"Significant" Estimates

A common method for decision-making under uncertainty is to accept as "significant" estimates whose margins of error are considered small, and using further decision rules for handling ambiguous cases not meeting a preset criterion for margin of error.

For example, a rule might establish one decision as significantly better than an alternative only if the estimated probability that the estimated decision is wrong is less than 5%. Where the estimated preference for a decision to medicate or not can be expressed along a continuum from strongly disfavour to strongly favour, the option to medicate based on an estimate of strongly favour is more likely to be *significantly* better than the option to not medicate, than where the estimated preference was for very slight leanings toward medication[20].

In most cases a multidimensional array of potential treatment options exists, in which many options can be individually varied as a continuous variable, and altered over time (e.g., hours of therapy, milligrams of Prozac). The average costs of a wrong decision with respect to continuous variable options is likely to be less than in a dichotomous situation where the two options are far apart, since in the dichotomous case a small error in preference estimation may result in a large decision error.

If, as has usually been the case, decisions to force treatment are characterized by decision-makers to be essentially dichotomous decisions — e.g., no drugs vs. the standard regimen — such decision processes can be expected to be, on average, less beneficial for the patient than decision processes that allow for marginal adjustments due to expressed patient preferences.

Treatment Reflexes

If the decision rule for ambiguous estimates is allowed to reflect non-patient interests — e.g., those of family or society, the kind of therapy the doctor enjoys doing, "follow the routine"[21] etc., then the pure best interests model will have been violated. Asserting that medical practitioners prefer to wrongly treat than to wrongly fail to treat, Scheff (1966) warned that a bias in favour of treatment may be a greater error than to not treat in the case of psychiatric care, because of the

---

[20] Note that the ordering and spacing of discrete option values on a continuum may be non-uniform and subjective (i.e., not a perfect Guttman scale; see Blalock Jr, 1979, pp. 22-23). For example, the secondary effects of a drug treatment might reach a watershed of greatly increased intolerability at a particular dosage.

[21] Denig, Haaijer-Ruskamp, Wesseling, and Versluis (1993); they report physician deviations from an expectancy-value model, which while saving time for the physician can be expected to detract from patient interests.

stigmatizing effects of diagnosis and treatment. Care must be taken to ensure that "short-cuts" are merely reasonable ways to simplify decision-making rather than decision rules imbued with non-patient values.

Bursztajn and Brodsky (1994), however, expressed concern about the danger of an increasing bias toward neglect, in an era of cost-cutting and managed care[22]. Whether viewed as matters of unjustifiable bias or as prudent decision rules in the case of uncertainty, any systematic reflex to treat or not to treat has to be carefully evaluated — particularly within the concrete environments that may reinforce such reflexes.

In the case of psychiatric care the threat to autonomy of a reflex to *treat* is obvious. Such a reflex involves relatively little expense, insofar as "treatment" implies drugs. It is less obvious, however, that there is a reflex to *care* — e.g. supportive services, housing and income — when it is relatively expensive yet crucial for supporting the development of patient autonomy. Despite frequently heard alarmist discourse about neglected drug-refusing psychiatric patients wandering the streets aimlessly, the alternative to "treatment" may be care, rather than neglect.

The leeway allowed by a rule incorporating an unevaluated reflex can, when combined with a conflict of interest, reduce the proxy's incentive to conscientiously estimate preferences and the margin of error in doing so. If drug therapy might serve the interests of a nursing home in the routine and orderly management of residents, for example, a proxy paid by that home might face coercion or subtle incentives to favour the choice of treatment by drugs on the grounds that despite evidence to the contrary the patient's *real* interests require drugs. The proxy could do this by expanding the perceived region of doubt —

---

[22] A bias toward neglect based on financial concerns but superficially justified in terms of patient autonomy worried Gunn (1991):

> The courts have decided to play some incompetent patient's hand for him without reference to the patient's actual wishes or of those in the best position to know those wishes. It was reasoned that if the patient were fully "rational," he would not want the treatment for himself .... A conscious competent patient's expressed wish to be kept alive will be ignored because his values do not coincide with the literature, the physician's beliefs, or societal needs. ... It would be embarrassing to present a rationale based on state or societal interest (pp. 149, 152).

giving the benefit of that doubt to the nursing home's interests[23]. Doubt about the patient's pure best interests does not excuse exploitation of that uncertainty by others with their own interests.

Criteria that the *Patient* Would Want

These types of difficulties might be reduced if criteria for close choices included rules like favouring the least restraint, the least physically invasive, the most reversible, giving the benefit of the doubt to expressed patient preference, etc. Such rules have to varying degrees guided court and policy decisions, and might be justifiable as consistent with the pure best interests philosophy as criteria that patients themselves might want for close choices.

Distortion of a rule has to be guarded against: e.g., where "least restraint *possible*" becomes "least restraint possible given the imposed treatment" or "danger to self or others" becomes in effect *risk to health*, meaning a mental illness which is not expected to improve without the imposed treatment (see Appelbaum, 1992; Cohen et al., 1997).

Decision Reversibility and Prospect of Future Competence. The reversibility of a decision is of crucial importance in the evaluation of treatment options, and has much to commend it as a means of deciding among close options. Decisions must be seen as part of a process over time, with each discrete decision altering the choices that will be available in the future (Arrow and Fisher, 1974; McCubbin, 1994a).

Hence, the more reversible a decision the greater the potential of improving overall outcomes through feedback and adjustment. This helps to explain people's preferences for risk that they can control, as discussed above. As observed by systems theorist Ackoff (1974), it is better to start with a poor solution that improves than with a good solution that gets worse.

---

[23] Regarding conflicts of interest that favour drug treatment see: Berlinguer, Falzi and Figà-Talamanca (1996); Cohen D. and McCubbin (1990); Doré and Cohen (1997); Lexchin (1993); Mohr (1994); Squires (1993).

A related consideration is the possibility that the patient will regain competence in the future. A substituted decision needs to both hasten and augment that eventuality as well as strive to not close off the options of the future competent patient (McCubbin, 1996). The implications of possible future competence for current substituted decisions have yet to be adequately explored.

Consequence of Error. Wrongly guessing that a candy is held in the right hand costs the same as wrongly guessing the left hand, hence the guesser will quickly make a "best guess" based only on the estimated probabilities that each hand holds the candy. A competent patient, acting as a rational decision maker with respect to important care and treatment decisions, will not make such a hasty best guess based only on probabilities; he or she will also assess the consequences of decision error taking into account the wide variety of direct and indirect effects of alternatives[24], which go well beyond the medical.

The preference estimation process must try to consider these consequences of error as the temporarily competent patient would have done. A process that does not do so will likely result in biased estimates. Chacko (1991) suggested that the severity of the two consequences of type 1 versus type 2 error should determine the null hypothesis; e.g., if convicting an innocent person is worse than freeing a guilty person, the null hypothesis becomes that the person is innocent.

The Cost of Forcing Treatment. However, consequence of error asymmetry arises not only in the calculation of the patient's preferences, but also in the imposition of a decision upon an unwilling patient. For example, the negative consequence of wrongly (according to "underlying" preferences) medicating when the patient says he or she does not want to be medicated is likely worse than for wrongly not medicating the same patient. The net benefits of medication will be greatly diminished when forced, due to the assault on the patient's sense of dignity and autonomy.

Imagine an incompetent patient who becomes competent and, knowing that the competence would continue, opts for treatment, knowing that each

---

[24] For discussions of balancing uncertain costs and benefits see: Mossman and Somoza (1992); Rossi and Freeman (1985, pp. 196-197).

administration of it will be consented to. Yet the same patient, if she knew she will immediately lapse into incompetence and knows that she would in the incompetent state vociferously object to the treatment might, consistent with her preferences, choose while momentarily competent not to receive the treatment. This would be due to the decision-maker taking account not only of the costs and benefits of the treatment itself but also of the psychic impact of having it forced — even though she knows that she would be better off not to refuse it while incompetent.

The medical decision facing a competent patient is only partly medical in its considerations; the decision of a proxy to force a medical intervention upon an unwilling patient is even less so[25]. The decision process must place weight upon the expressed preference not only as an indication of "underlying" preferences, but because the error of forcing an unwanted intervention is worse than the error of deciding in favour of what the patient requests.

Therefore, in addition to acting on the basis of what the patient "really" wants, an additional degree of certainty must be obtained before forcing treatment upon an unwilling patient. A good proxy decision process on behalf of patients who express preferences must be prepared to wrongly reject some estimated

---

[25] A competent patient's "medical" decision becomes the proxy's *metamedical* decision, in that it must take into account the impact of an imposed proxy process upon the patient. General systems theory (GST) can aid in contextualizing medical intervention:

> A systems view confronts us with the awareness that although therapeutic interventions are primarily aimed at one level of the system, they will indirectly affect other levels as well; furthermore, homeostasis reestablished at one level may be purchased at the cost of greater dysfunction at other levels. The contribution of this model to ethical considerations in therapeutics rests in its ability to help us see more clearly and with less distortion the full range of impact of therapy on all levels of the organism ... (Sider 1984, p. 392)

For surveys of the application of GST within the mental health field, particularly at the policy level, see McCubbin (1994b) and McCubbin and Cohen (1997). Engel (1980) drew heavily on GST in developing his biopsychosocial model of clinical care. Bateson's (1972) psychology and efforts to grasp an understanding of schizophrenia is imbued with GST philosophy. Bernstein and Lennard (1973) modeled the systemic context and interactions of psychiatric drug treatment. Ziegenfuss Jr (1983) applied a variant of GST, sociotechnical systems theory, to the problem of creating patient rights programs in an institution. Blum (1983) described how a national health policy could be developed according to GST principles.

preferences in order to accord with expressed preferences, insofar as the detrimental effect of forcing a decision outweighs the net benefits of that decision according to the estimated preferences.

Reasonable Certainty. One way to view the problem of choosing among uncertain alternatives is to consider one option the "null hypothesis", with alternatives having to be "proved". A rule of thumb in the social sciences is to accept an alternative as probably true if there is less than a 5% chance of being wrong, and otherwise we rest with the null hypothesis. This level of probability might be called *reasonable certainty* which, while in theory perhaps less certain than "beyond a reasonable doubt", is more certain than that implied by "preponderance of evidence".

An *inertial principle* would be to define the null hypothesis as inaction, the status quo, or non-intervention; e.g., the proxy would have to be 95% sure of being right before approving drug treatment. While this approach might be seen as reasonable in that it will often serve the purpose of other justifiable criteria (e.g., non-intervention might ordinarily be the patient's expressed preference in mandated cases, and might also ordinarily be the least intrusive and most reversible), it is not clear that this criterion would be right for all situations, and furthermore, it could be replaced by the criteria used to justify it. It does, however, have the advantage of simplicity.

Given the costs of forced treatment discussed above, it seems reasonable to adopt the patient's expressed preference as the null hypothesis, on the ethical ground that *very* strong justification has to be advanced before impairing the patient's freedom of choice. That it be *reasonably certain* that the decision is correct should then be considered, in addition to demonstrated incompetence, a necessary condition for making a treatment decision against the patient's expressed will.

The Implications of Uncertainty for Forced Treatment

The difficulty of validly ascertaining a patient's pure best interests and accordingly making a reliable and unbiased substituted decision, a decision which appropriately weights the costs of error and the impact of forcing treatment, might suggest severe limitations in the justifiability of mandated treatment. Two kinds of patient incompetence in particular might justify such substituted decision-making: 1) where the patient is unable to *express* preferences; 2) where the patient is unable to give informed consent because it is demonstrably clear that the patient cannot comprehend relevant information such as what the decision options are, their possible consequences, and the risks of their consequences.

However, even if a patient is clearly incapable of making an informed treatment decision, it does not follow that one should be imposed. This article argues that a further necessary condition for the imposition of treatment is reasonably certainty that the patient's underlying preferences — including those bearing on the subjective impacts of having a treatment imposed — have been correctly determined[26]. Not only the patient's estimated treatment outcome preferences, but also the patient's attitude to risk and personal discount rate have to be considered.

These factors combine to create a region of uncertainty, measured by the estimated error variance of the proxy decision-making process, that will properly mean that some patients will not have a treatment imposed when they would have preferred that it would be. Accepting this cost avoids a greater cost where patients are coercively treated who wouldn't want to be (see Einhorn, 1988).

It needs to be noted that the "best" decision is rarely the most uncertain or risk-free. Taking risks and making mistakes can be a therapeutic learning experience (Ryan 1993), when based on unfettered conation. Obviously, the degree of uncertainty can be reduced in most cases by conducting a more exhaustive research of the patient as a unique contextualized individual. The science and art of estimating preferences is barely nascent and requires much growth in terms of how to assess the reliability, weight and bias to apply to each source of information, to develop indicators, and to make good decisions given the uncertainty that will remain.

It is important to bear in mind that there is much controversy over the objectivity and reliability of incompetence determinations (Winick 1996). The

---

[26] Note that nothing in this article implies that incompetence and reasonable certainty provide a *sufficient* justification for forced treatment. Indeed, my view is that they do *not*, in the case of psychiatric care, for reasons which go beyond the purview of this article.

probability of wrongly forcing treatment is augmented not only by risks of decision alternatives, error in estimating preferences, and failure to account for the impact on the patient of forcing treatment, but also by the possibility of error in the decision to declare the patient incompetent.

My own view is that the synthesis of these four sources of error — too often systematically related and reinforcing — result in far more cases of forced treatment upon objecting (or passive) psychiatric patients than would be justifiable under adequately conducted and evaluated pure best interests procedures.

Complexity

A major difficulty with pure best interests substituted decision-making lies in the ability of a proxy decision maker to effectively carry out the decision-making process itself. The competent patient's process of determining his or her own preferences can be very complex, incorporating as it does probability distributions of various possible outcomes and the likely interaction of various variables in the utility function (e.g., negative utility to separate consumption of coffee and cigarettes, positive utility to joint consumption).

The proxy decision maker has an even more complex task, having to guess at, rather than know, likes and dislikes, incorporate in the process a measure of uncertainty that the estimated preferences are wrong, and finally to select and apply appropriate decision rules given decision reversibility, a possible expectation of the patient regaining capacity, the distribution of costs and benefits for decision alternatives, the patient's risk utility function, the patient's personal discount rate, and the impact of forcing treatment. The relations or functions involved are unlikely to lend themselves to simple analysis: variables might be only multinomial rather than ratio-level, distributions might be asymmetric, etc.[27]

In modeling the decision problem of setting dangerousness criteria, using violence predictions and cost-benefit measures, Mossman and Somoza (1992) observed that error in utility and probability estimates will "greatly affect the

---

[27] See: Goldman, Rachuba, and Van Tosh (1995); Loehman (1994); Pellissier and Hazen (1994).

number of nonviolent persons who are hospitalized", yet "risk-benefit balancing is a strategy whose application has severe mathematical limitations" (p. 335). Proxy decision-making can never be a perfectly reliable scientific activity. Knowing this does not, however, excuse lax procedures that have no scientific basis or ethical justifiability. Rather, it is incumbent upon us to do the best we can, within reason, and be honest about the weaknesses of our procedures and the implications of that weakness for how and what we decide.

Due to the increased complexity involved in a well-made proxy decision-making process, there is a danger that a proxy will reduce the level of information processing below what would be expected, as a result of "information overload and ensuing fatigue" (Janis and Mann, 1977, p. 17). But in making a reasonable choice on behalf of others rather than for one's self the decision process needs to become less intuitive and more explicit (if not, the choice is likely to be very biased toward the proxy's values) as well as more time consuming. Without supervision, there may be reason to fear that proxies will tend to use more short-cuts in the decision-making process than is justifiable, given that the incentives to make a good decision are not the same for the proxy as they would be for a competent patient.

Monitoring Proxy Processes

Hence the burden of proxy decision-making should not rest entirely upon one person's shoulders — particularly if that person does not have the training, capacity or resources to adequately conduct the preference estimation process. The skills required, particularly in complex cases with heavy consequences for the patient — e.g. when considering forced treatment — go beyond those typically held by a loving family member.

There is also no reason to think that the skill of estimating preferences and making a substituted decision accordingly is a medical skill; hence the participation of fellow patients and ex-patients, advocates, courts and social workers in the decision-making process, perhaps aided by ethicists or decision analysts, can be justified insofar as they are given adequate resources and are trained in the particular skills required to ascertain and protect the pure best interests of those they aim to help.

Conclusion: Good Intentions Are Not Good Enough

While the issue of involuntary psychiatric treatment has, for several decades, been one of the most discussed in the literature focused on the mental health system, an important link in the argument for forced treatment has barely been tested. Controversy over the state of psychiatric knowledge, debates over the relative efficacies of competing approaches to mental health, ethical arguments over civil rights and due process, and often emotional critique, denial or defense of the personal, political and institutional power of psychiatry, have not led attention to a necessary prerequisite for imposition of psychiatric treatment against the will of a patient: whether, even if we were to believe that a patient is "not in his or her right mind", it is possible to determine that person's own "best interests" for the purpose of substituted decision-making; and if so, how, to what degree, and under what conditions.

Whether a patient should be allowed to choose outcomes for him or herself or whether choices should be made by others on behalf of the patient depends not only on establishing the patient's incompetence, but also on *the degree to which the patient's preferences can and will be estimated and followed.* Whether psychiatrists, family members, and others usually involved in proxy decisions have the incentives to conscientiously try to ascertain and follow patient preferences is problematic, but not dealt with here (see McCubbin and Cohen D. 1996). This article raises the complicated considerations and problems entailed in determining patient preferences for the purpose of proxy decision-making under the pure best interests model, by discussing the sources of uncertainty involved in the preference estimation process, and ways in which to deal with or acknowledge them.

Where a competent patient makes her/his own treatment decisions, risk enters only in terms of the uncertain nature of potential outcomes. Estimates of medical risk factors are ordinarily expected to be supplied by treating professionals; as Gigliotti and Rubin (1991) pointed out, however, in evaluating the merits of the uncertain choices the patient then applies not only valuations of outcomes, but also an attitude to risk.

However, a process of substituted decision-making adds a much greater degree of uncertainty: even careful efforts to estimate the patient's preferences do not guarantee that a reasonably valid, unbiased or reliable substituted treatment decision will be arrived at. Validity problems arise where a proxy decision maker fails to recognize the subjective nature of assessing, valuing and controlling alternative outcome risks, and where the proxy constructs preference estimates around a former or future patient rather than around the present individual.

Bias resulting from participants or sources with interests, values or perspectives at odds with the patients' is an important problem which might be reduced with extensive research, by identifying and compensating for the interests of sources, and especially by ensuring the independence of the proxy from interest in the decision outcomes other than in satisfying the patient's pure best interests. This will not be the case where a psychiatrist diagnoses a mental illness, assesses ensuing incompetency, and decides upon treatment.

The discussion of estimation error variance raises the very difficult problem of estimating the expected error distribution. There is no good way to obtain historical data on the degree to which estimates varied from preferences in previous cases, particularly given the dubious scientific value of a post treatment "thankyou".

The problem of how to make a decision in the face of uncertainty has been discussed at length. How uncertainty is resolved must be justifiable in terms of the patient's pure best interests; benefit of the doubt might accordingly be given to options that are more reversible, less invasive or restraining, and which accord with the patient's expressed preferences. This article suggests that due to the impacts of forcing a treatment upon an unwilling patient, it should be *reasonably certain* that the treatment meets the patient's underlying preferences.

It bears repeating that estimating the preferences of others is — or should be — far more difficult than determining one's own preferences: the human brain ordinarily has the capacity to simultaneously incorporate many factors with probabilities and values to reach a decision fairly quickly even if not infallibly; to do so reasonably well on someone else's behalf, however, requires an explicit and careful process in order to cope with the additional complexity and uncertainty. Further research is needed to develop proxy decision-making models that incorporate a sense of their own fallibility. Furthermore, such models need to be compared with actual proxy procedures in their diverse contexts, as part of a dialectical process of creating useful theory and ethical practice.

"Feasibility" as an Ethical Question

A reader of a much earlier version of this article made the very pertinent comment that while it provides a neat way to analyze the problem of substituted decision-making and point the way to further research, the feasibility of the "advocated cost/benefit analysis" is open to question. The motivation behind this exposition, however, was not so much to establish a new method for substituted decision-making, but rather to stimulate discussion of the conditions required before society could ethically justify imposing treatment upon persons diagnosed with mental illness.

Indeed, the question of feasibility is crucial to the ethical justifiability of substituted decision-making. At least in the case of patients able to express their own wishes, *if proxy decision-making cannot be done reliably and well, it should not be done at all.* Hence, if it seems not feasible to make a reasonable effort to determine the pure best interests of a patient for whom forced treatment is being considered, the alternative is *not* to substitute a much easier method that has no clear objectives, cannot be evaluated, and enables the camouflaged entry of non-patient interests. Rather, the alternative is to admit that we have no business in interfering in the autonomy of patients who express their own preferences, because we would have no justification to do so.

Any time society, or those with exceptional power within society, imposes unwanted psychiatric treatment on an individual, there is an obligation to demonstrate with at least reasonable certainty that such an imposed treatment, as well as the fact of *imposing* it, really serves the interests of the patient rather than the interests of others. It is partly for this reason — to avoid the paternalism inherent in the best interests model that can hide non-patient interests — that the *pure* best interests model has been favoured here as ethically preferable if substituted decision-making seems to become necessary.

Today, however, many individuals have their liberties removed due to diagnoses of mental illness and, in the discourse of more "advanced" jurisdictions, due to incompetence ensuing from those illnesses. Psychiatric treatments, consisting today almost exclusively of emotion and mind altering drugs, are imposed upon those individuals even though the efficacy of the treatments are usually dubious, especially over the long run, most incur uncomfortable or disabling secondary effects, insufficient attention has been given to more empowering and less physically intrusive help such as psychosocial care, income, housing and community support and, as this paper emphasizes, *with no good reason to believe that those who make the decisions for those patients ascertain their pure best interests with a reasonable degree of accuracy and reliability.* Insofar as current practices in the mental health system are at the limits of their feasibility, we cannot justify imposing psychiatric decisions upon unwilling patients.

References

Ackoff, R. L. (1974). The systems revolution. *Long Range Planning, 7,* 2-20.

Anderson, J., and Eppard, J. (1995). Clinical decision making during assessment for involuntary psychiatric admission. *Psychiatric Services, 46,* 727-728.

Antonovsky, A. (1993). Complexity, conflict, chaos, coherence, coercion and civility. *Social Science and Medicine, 37,* 969-981.

Appelbaum, P. S. (1992). Legal aspects of clinical care for severely mentally ill, homeless persons. *Bulletin of the American Academy of Psychiatry and Law, 20,* 455-473.

Appelbaum, P. S., and Grisso, T. (1995). The MacArthur Treatment Competence Study. I: Mental illness and competence to consent to treatment. *Law and Human Behavior, 19,* 105-126.

Appelbaum, P. S., and Schwartz, W. F. (1992). Minimizing the social cost of choosing treatment for the involuntarily hospitalized mentally-ill patient: A new approach to defining the patient's role. *Connecticut Law Review, 24,* 433-485.

Arrow, K. J., and Fisher, A. C. (1974). Environmental preservation, uncertainty, and irreversibility. *Quarterly Journal of Economics, 88,* 312-319.

Baergen, R. (1995a). Revising the substituted judgment standard. *Journal of Clinical Ethics, 6,* 30-38.

Baergen, R. (1995b). Surrogates and uncertainty. *Journal of Clinical Ethics, 6,* 372-377.

Bassford, H. A. (1990). The basis of medical ethics. In D. MacNiven (Ed.), *Moral expertise: Studies in practical and professional ethics* (pp. 128-143). New York: Routledge.

Bateson, G. (1972). *Steps to an ecology of mind.* New York: Ballantine.

Becker, G., and Nachtigall, R. D. (1994). 'Born to be a mother': The cultural construction of risk in infertility treatment in the U.S. *Social Science and Medicine, 39,* 507-518.

Berlinguer, G., Falzi, G., and Figà-Talamanca, I. (1996). Ethical problems in the relationship between health and work. *International Journal of Health Services, 26,* 147-171.

Bernstein, A., and Lennard, H. L. (1973, May/June). The American way of drugging: Drugs, doctors and junkies. *Society,* 14-25.

Blalock, H. M. Jr (1979). *Social statistics* (Revised Second Edition). New York: McGraw-Hill.

Blum, H. L. (1983). *Expanding health care horizons: From a general systems concept of health to a national health policy* (2nd edition). Oakland, CA: Third Party.

Borges, E. M. (1995). A social critique of biological psychiatry. In C. A. Ross and A. Pam (Eds.), *Pseudoscience in biological psychiatry: Blaming the body* (pp. 211-240). Toronto: Wiley.

Brady, M. E. (1993). J. M. Keynes's theoretical approach to decision-making under conditions of risk and uncertainty. *British Journal for the Philosophy of Science, 44,* 357-376.

Brock, D. W. (1994). Good decision-making for incompetent patients. *Hastings Center Report, 24*(6), S8-S11.

Bursztajn, H. J., and Brodsky, A. (1994). Authenticity and autonomy in the managed-care era: Forensic psychiatric perspectives. *Journal of Clinical Ethics, 5,* 237-242.

Campinha-Bacote, J. (1994). Transcultural psychiatric nursing: Diagnostic and treatment issues. *Journal of Psychosocial Nursing, 32,* 41-46.

Carmines, E. G., and Zeller, R. A. (1986). *Reliability and validity assessment.* Beverly Hills: Sage.

Carpentier-Roy, M. C. (1995). Anomie sociale et recrudescence des problèmes de santé mentale au travail. *Santé mentale au Québec, 20,* 119-138.

Chacko, G. K. (1991). *Decision-making under uncertainty: An applied statistics approach.* New York: Praeger.

Chewning, B., and Sleath, B. (1996). Medication decision-making and management: A client-centered model. *Social Science and Medicine, 42,* 389-398.

Christensen-Szalanski, J. J. J., and Bushyhead, J. B. (1988). Physicians' use of probabilistic information in a real clinical setting. In J. Dowie and A. Elstein (Eds.), *Professional judgment: A reader in clinical decision making* (pp. 360-373). Cambridge: Cambridge University Press.

Cohen, C. I. (1993). The biomedicalization of psychiatry: A critical overview. *Community Mental Health Journal, 29,* 509-521.

Cohen, D., and McCubbin, M. (1990). The political economy of tardive dyskinesia: Asymmetries in power and responsibility. *Journal of Mind and Behavior, 11,* 465-488.

Cohen, D., Thomas, G., Dallaire, B., McCubbin, M., Morin, P., and Fortier, R. (1997). Savoir, pouvoir, et dangérosité civile: Une étude quantitative et qualitative des décisions de révision de cure fermée de la Commission des affaires sociales du Québec, 1975-1993. Unpublished.

Cornes, R., and Sandler, T. (1986). *The theory of externalities, public goods, and club goods.* Cambridge: Cambridge University Press.

Crystal, S., and Dejowski, E. (1987). Substituted judgment and protective intervention. *New Directions for Mental Health Services, 36,* 83-91.

Denig, P., Haaijer-Ruskamp, F. M., Wesseling, H., and Versluis, A. (1993). Towards understanding treatment preferences of hospital physicians. *Social Science and Medicine, 36,* 915-924.

Doré, C., and Cohen, D. (1997). La prescription de stimulants aux enfants "hyperactifs": Une étude pilote des incitatifs et des contraintes pour les parents, les médecins et les enseignants. *Santé mentale au Québec, 22,* 216-238.

Druley, J. A., Ditto, P. H., Moore, K. A., Danks, J. H., Townsend, A., and Smucker, W. D. (1993). Physicians' predictions of elderly outpatients' preferences for life-sustaining treatment. *Journal of Family Practice, 37,* 469-475.

Einhorn, H. J. (1988). Accepting error to make less error. In J. Dowie and A. Elstein (Eds.), *Professional judgment: A reader in clinical decision making* (pp. 181-189). Cambridge: Cambridge University Press.

Ellis, R. D. (1993). Quantifying distributive justice: An approach to environmental and risk-related public policy. *Policy Sciences, 26,* 99-123.

Engel, G. L. (1980). The clinical application of the biopsychosocial model. *American Journal of Psychiatry, 137,* 555-564.

Eppard, J., and Anderson, J. (1995). Emergency psychiatric assessment: The nurse, psychiatrist and counselor roles during the process. *Journal of Psychosocial Nursing, 33*(10), 17-23.

Fabrega, H. Jr (1992). The role of culture in a theory of psychiatric illness. *Social Science and Medicine, 35,* 91-103.

Fischoff, B., and Beyth-Marom, R. (1988). Hypothesis evaluation from a Bayesian perspective. In J. Dowie and A. Elstein (Eds.), *Professional judgment: A reader in clinical decision making* (pp. 323-348). Cambridge: Cambridge University Press.

Frank, R. H. (1985). *Choosing the right pond: Human behavior and the quest for status.* New York: Oxford University Press.

Freidson, E. (1984). La profession médicale. Paris: Poyot.

Frey, B. S. (1994). How intrinsic motivation is crowded out and in. *Rationality and Society, 6,* 334-352.

Furnham, A. (1994). Explaining health and illness: Lay perceptions on current and future health, the causes of illness, and the nature of recovery. *Social Science and Medicine, 39,* 715-725.

Gerrity, M. S., Earp, J. A. L., DeVellis, R. F., and Light, D. W. (1992). Uncertainty and professional work: Perceptions of physicians in clinical practice. *American Journal of Sociology, 97,* 1022-1051.

Gigliotti, G. A., and Rubin, J. (1991). The right to refuse treatment: An application of the economic principles of decision-making under uncertainty. *International Journal of Law and Psychiatry, 14,* 405-416.

Goldman, H. H., Rachuba, L., and Van Tosh, L. (1995). Methods of assessing mental health consumers' preferences for housing and support services. *Psychiatric Services, 46,* 169-172.

Gordon, R. M. (1993). Out to pasture: A case for the retirement of Canadian mental health legislation. *Canadian Journal of Community Mental Health, 12,* 37-55.

Greenberg, A. S., and Bailey, J. M. (1994). The irrelevance of the medical model of mental illness to law and ethics. *International Journal of Law and Psychiatry, 17,* 153-173.

Griffith, D. B. (1991). The best interests standard: A comparison of the state's *parens patriae* authority and judicial oversight in best interests determinations for children and incompetent patients. *Issues in Law & Medicine, 7,* 283-338.

Grisso, T., Appelbaum, P. S., Mulvey, E. P., and Fletcher, K. (1995). The MacArthur Treatment Competence Study. II: Measures of abilities related to competence to consent to treatment. *Law and Human Behavior, 19,* 127-148.

Gunn, A. E. (1991). Risk-benefit ratio: The soft underbelly of patient autonomy. *Issues in Law & Medicine, 7,* 139-153.

Gustafson, D. H., Sainfort, F., Johnson, S. W., and Sateia, M. (1993). Measuring quality of care in psychiatric emergencies: Construction and evaluation of a Bayesian index. *Health Services Research, 28,* 131-158.

Haas, J., and Shaffir, W. (1987). Taking on the role of doctor: A dramaturgic analysis of professionalization. In D. Coburn, C. D'Arcy, G. M. Torrance, and P. New (Eds), *Health and Canadian society: Sociological perspectives* (Second Edition) [pp. 399-421]. Markham, Ontario: Fitzhenry & Whiteside.

Hammitt, J. K. (1993). Discounting health increments. *Journal of Health Economics, 12,* 117-120.

Hayes, M. V. (1992). On the epistemology of risk: Language, logic and social science. *Social Science and Medicine, 35,* 401-407.

Hermann, D. H. J. (1990). Autonomy, self determination, the right of involuntarily committed persons to refuse treatment, and the use of substituted judgment in medication decisions involving incompetent persons. *International Journal of Law and Psychiatry, 13,* 361-385.

Hey, J. D., and Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62,* 1291-1326.

Hilden, J., Glasziou, P. P., and Habbema, J. D. F. (1992). A pitfall in utility assessment: Patients' undisclosed investment decisions. *Medical Decision Making, 12,* 39-43.

Howson, C., and Urbach, P. (1991). Bayesian reasoning in science. *Nature, 350,* 371-374.

Janis, I. L., and Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.

Kaufman, S. R. (1995). Decision making, responsibility, and advocacy in geriatric medicine: Physician dilemmas with elderly in the community. *Gerontologist, 35,* 481-488.

Klein, W. M., and Kunda, Z. (1994). Exaggerated self-assessments and the preference for controllable risks. *Organizational Behavior and Human Decision Processes, 59,* 410-427.

Kline, S. A. (1992). Substitute decision-making and the substitute decision-maker. *Health Law in Canada, 13,* 125-128.

Koffman, J., Fulop, N. J., Pashley, D., and Coleman, K. (1997). Ethnicity and use of acute psychiatric beds: One-day survey in North and South Thames regions. *British Journal of Psychiatry, 171,* 238-241.

Kopelman, L. M. (1990). On the evaluative nature of competency and capacity judgments. *International Journal of Law and Psychiatry, 13,* 309-329.

Laor, N. (1984). Common sense ethics and psychiatry. *Psychiatry, 47,* 135-150.

Lee, S. (1994). *Heller v. Doe*: involuntary civil commitment and the "objective" language of probability. *American Journal of Law and Medicine, 20,* 457-477.

Lesemann, F., and Chaume, C. (1990). Modes de prise en charge familiaux: Quand les pratiques de maintien à domicile interrogent le bien-fondé des politiques. *Gérontophile, 12*(4), 19-22.

Lexchin, J. (1993). Interactions between physicians and the pharmaceutical industry: What does the literature say? *Canadian Medical Association Journal, 149,* 1401-1407.

Littlewood, R. (1992). Psychiatric diagnosis and racial bias: Empirical and interpretative approaches. *Social Science and Medicine, 34,* 141-149.

Loehman, E. (1994). Rank dependent expected utility: Stochastic dominance, risk preference, and certainty equivalence. *Journal of Mathematical Psychology, 38,* 159-197.

Loewenstein, G. F., and Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review, 100,* 91-108.

Loring, M., and Powell, B. (1988). Gender, race, and DSM-III: A study of the objectivity of psychiatric diagnostic behavior. *Journal of Health and Social Behavior, 29,* 1-22.

Lupton, D. (1993). Risk as moral danger: The social and political functions of risk discourse in public health. *International Journal of Health Services, 23,* 425-435.

Lynn, J. (1992). Procedures for making medical decisions for incompetent adults. *JAMA, 267,* 2082-2084.

Lynn, J., and Teno, J. M. (1993). After the Patient Self-Determination Act: The need for empirical research on formal advance directives. *Hastings Center Report, 23*(1), 20-24.

Martin, S. C., Arnold, R. M., and Parker, R. M. (1988). Gender and medical socialization. *Journal of Health and Social Behavior, 29,* 333-343.

McCubbin, M. (1994a). Deinstitutionalization: The illusion of disillusion. *Journal of Mind and Behavior, 15,* 35-53.

McCubbin, M. (1994b). *A systems theory perspective of mental health services and policy formation* (examen de synthèse). Montréal: Université de Montréal.

McCubbin, M. (1996). How the "right" decision to force psychotropic drug treatment can be wrong: Consequence of error asymmetry. Unpublished.

McCubbin, M. (1997). Population health: A call for breadth (mental health) and depth (psychosocial theory). Submitted to *Social Science and Medicine.*

McCubbin, M., and Cohen, D. (1996). Extremely unbalanced: Interest divergence and power disparities between clients and psychiatry. *International Journal of Law and Psychiatry, 19,* 1-25.

McCubbin, M., and Cohen, D. (1997). Mental health policy from a systemic perspective: Orientations for strategic reform. Submitted to *Journal of Health Politics, Policy and Law.*

McCubbin, M., and Weisstub, D. (1998). Toward a pure best interests model of proxy decision-making for incompetent psychiatric patients. In press, *International Journal of Law and Psychiatry, 21*(1).

McEvoy, J. P., Aland, J. Jr, Wilson, W. H., Guy, W., and Hawkins, L. (1981). Measuring chronic schizophrenic patients' attitudes toward their illness and treatment. *Hospital and Community Psychiatry, 32,* 856-858.

McKenzie, R. B. (1991). Rational addiction, lagged demands and the efficiency of excise taxes: Revisions of standard theory. *Public Choice, 71,* 33-41.

Meisel, A. (1979). The "exceptions" to the informed consent doctrine: Striking a balance between competing values in medical decisionmaking. *Wisconsin Law Review, 127,* 413-488.

Miedema, B. (1994). Control or treatment? Experiences of people who have been psychiatrically hospitalized in New Brunswick. *Canadian Journal of Community Mental Health, 13,* 111-122.

Mitchell, J. (1990). The evolution of acceptable risk. *The Environmental Professional, 12,* 114-121.

Mohr, W. K. (1994). The private psychiatric hospital scandal: A critical social approach. *Archives of Psychiatric Nursing, 7,* 3-8.

Mohr, W. K. (1995). Values, ideologies, and dilemmas: Professional and occupational contradictions. *Journal of Psychosocial Nursing, 33,* 29-34.

Mossman, D., and Somoza, E. (1992). Balancing risks and benefits: Another approach to optimizing diagnostic tests. *Journal of Neuropsychiatry and Clinical Neurosciences, 4,* 331-335.

Nease, R. F., Kneeland, T., O'Connor, G. T., Sumner, W., Lumpkins, C., Shaw, L., Pryor, D., and Sox, H. C. (1995). Variation in patient utilities for outcomes of the management of chronic stable angina: Implications for clinical practice guidelines. *JAMA, 273,* 1185-1190.

Olsen, J. A. (1993). On what basis should health be discounted? *Journal of Health Economics, 12,* 39-53.

Parry, J. (1987). A unified theory of substitute consent: Incompetent patients' right to individualized health care decision-making. *Mental and Physical Disability Law Reporter, 11,* 378-441.

Pellissier, J. M., and Hazen, G. B. (1994). Implementation of continuous risk utility assessment: The total hip replacement decision. *Socio-Economic Planning Sciences, 28,* 251-276.

Pellissier, J. M., Hazen, G. B., and Chang, R. W. (1996). A continuous-risk decision analysis of total hip replacement. *Journal of the Operational Research Society, 47,* 776-793.

Perlin, M. L. (1990). Are courts competent to decide competency questions?: Stripping the facade from *United States v. Charters. Kansas Law Review, 38,* 957-1001.

Popper, K. R. (1979). *Objective knowledge: An evolutionary approach.* Oxford: Oxford University Press.

Pruchno, R. A., Smyer, M. A., Rose, M. S., Hartman-Stein, P. E., and Henderson-Laribee, D. L. (1995). Competence of long-term care residents to participate in decisions about their medical care: A brief, objective assessment. *Gerontologist, 35,* 622-629.

Rabinowitz, J. (1993). Diagnostic reasoning and reliability: A review of the literature and a model of decision-making. *Journal of Mind and Behavior, 14,* 297-316.

Redding, R. E. (1993). Children's competence to provide informed consent for mental health treatment. *Washington and Lee Law Review, 50,* 695-753.

Redelmeier, D. A., and Heller, D. N. (1993). Time preference in medical decision making and cost-effectiveness analysis. *Medical Decision Making, 13,* 212-217.

Reich, W. (1984). Psychiatric diagnosis as an ethical problem. In S. Bloch and P. Chodoff (Eds.), *Psychiatric ethics* (paperback corrected version) [pp. 61-88]. Oxford: Oxford University Press.

Rifkin, W. D. (1994). Who need not be heard: Deciding who is not an expert. *Technology Studies, 1,* 60-96.

Robertson, J. A. (1985). The geography of competency. *Social Research, 52,* 555-579.

Rodwin, M. A. (1994). Patient accountability and quality of care: Lessons from medical consumerism and the patients' rights, women's health and disability rights movements. *American Journal of Law and Medicine, 20,* 147-167.

Rogler, L. H. (1993). Culturally sensitizing psychiatric diagnosis: A framework for research. *Journal of Nervous and Mental Disease, 181,* 401-408.

Ronis, D. L. (1992). Conditional health threats: Health beliefs, decisions, and behaviors among adults. *Health Psychology, 11,* 127-134.

Rosenfield, S. (1992). Factors contributing to the subjective quality of life of the chronic mentally ill. *Journal of Health and Social Behavior, 33,* 299-315.

Rossi, P. H., and Freeman, H. E. (1985). *Evaluation: A systematic approach* (Third Edition). Beverly Hills: Sage.

Ryan, T. (1993). Therapeutic risks in mental health nursing. *Nursing Standard, 7*(24), 29-31.

Scheff, T. J. (1966). *Being mentally ill: A sociological theory*. Chicago: Aldine.

Scheid, T. L. (1994). An explication of treatment ideology among mental health care providers. *Sociology of Health & Illness, 16*, 668-693.

Scott, A., Shiell, A., and King, M. (1996). Is general practitioner decision making associated with patient socio-economic status? *Social Science and Medicine, 42*, 35-46.

Sider, R. C. (1984). The ethics of therapeutic modality choice. *American Journal of Psychiatry, 141*, 390-394.

Smithson, M. (1987). *Fuzzy set analysis for behavioral and social sciences*. New York: Springer-Verlag.

Sorum, P. C. (1995). Deciding about cardiopulmonary resuscitation: The contributions of decision analysis. *Archives of Internal Medicine, 155*, 513-521.

Squires, B. P. (1993). Physicians and the pharmaceutical industry. *Canadian Medical Association Journal, 149*, 1391-1392.

Stambolovic´, V. (1996). Human rights and health within the dominant paradigm. *Social Science and Medicine, 42*, 301-303.

Stein, H. F. (1990). *American medicine as culture*. Boulder, CO: Westview.

Strudler, A. (1988). Self-determination, incompetence, and medical jurisprudence. *Journal of Medicine and Philosophy, 13*, 349-365.

Sullivan, W. P. (1992). Reclaiming the community: The strengths perspective and deinstitutionalization. *Social Work, 37*, 204-209.

Sulmasy, D. P., Haller, K., and Terry, P. B. (1994). More talk, less paper: Predicting the accuracy of substituted judgments. *American Journal of Medicine, 96*, 432-438.

Syx, C. (1995). The mental health service system: How we've created a make-believe world. *Psychiatric Rehabilitation Journal, 19*, 83-85.

Tomossy, G. F., and Weisstub, D. N. (1997). The reform of adult guardianship laws: The case of non-therapeutic experimentation. *International Journal of Law and Psychiatry, 20*, 113-139.

Tsevat, J., Cook, E. F., Green, M. L., Matchar, D. B., Dawson, N. V., Broste, S. K., Wu, A. W., Phillips, R. S., Oye, R. K., and Goldman, L. (1995). Health values of the seriously ill. *Annals of Internal Medicine 122*, 514-520.

Vatz, R. E., and Weinberg, L. S. (1990). The conceptual bind in defining the volitional component of alcoholism: Consequences for public policy and scientific research. *Journal of Mind and Behavior, 11,* 531-544.

Venesy, B. A. (1994). A clinician's guide to decision making capacity and ethically sound medical decisions. *American Journal of Physical Medicine & Rehabilitation, 73,* 219-226.

Wade, J. C. (1993). Institutional racism: An analysis of the mental health system. *American Journal of Orthopsychiatry, 63,* 536-544.

Winick, B. J. (Ed.). (1996). A critical examination of the MacArthur Treatment Competence Study: Methodological issues, legal implications, and future directions. *Psychology, Public Policy, and Law,* 2(1).

Winick, B. J. (1997). *The right to refuse mental health treatment.* Washington, DC: American Psychological Association.

Wolf, F. M., Gruppen, L. D., and Billi, J. E. (1988). Differential diagnosis and the competing-hypotheses heuristic: A practical approach to judgment under uncertainty and Bayesian probability. In J. Dowie and A. Elstein A. (Eds.), *Professional judgment: A reader in clinical decision making* (pp. 349-359). Cambridge: Cambridge University Press.

Wolf-Branigin, M., and Sawilowsky, S. (1994). Perceived importance of self-determination factors by consumers in substance-abuse treatment. *Perceptual and Motor Skills, 79,* 284-286.

Wolff, N. (1989). Professional uncertainty and physician medical decision-making in a multiple treatment framework. *Social Science and Medicine, 28,* 99-107.

Wulff, H. R. (1981). How to make the best decision: Philosophical aspects of clinical decision theory. *Medical Decision Making, 1,* 277-283.

Ziegenfuss, J. T. Jr (1983). *Patients' rights and organizational models: Sociotechnical systems research on mental health programs*. Washington, DC: University Press of America.

"Meeting the Needs of the Mentally Ill"

A Case Study of the "Right to Treatment" as Legal Rights Discourse in the U.S.A.

Michael McCubbin* and David N. Weisstub**

*Groupe de recherche sur les aspects sociaux de la santé et de la prévention (GRASP), Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec), Canada, H3C 3J7

[redacted]

**Chaire de psychiatrie légale et d'éthique biomédicale Philippe Pinel, Faculté de médecine, Université de Montréal

Acknowledgments—We appreciate the thoughtful comments on earlier drafts from David Cohen, GRASP, David Rochefort, Public Administration Program, Northeastern University, and two anonymous readers. This paper will be further revised, further taking into account those comments, aiming to submit it to a journal as a co-authored article. Preparation of this article was partly funded by the Québec Réseau d'éthique clinique chez l'humain (Unité psychiatrie et psychologie). Michael McCubbin has received generous financial support over the last several years from the Université de Montréal, through bursaries and research contracts.

6 February 1998

Abstract—We examine the history of one of the earliest bioethics issues in the United States, the "right to treatment", and the changes in how this issue has been framed in recent years. Right to treatment litigation raised awareness of other "patient rights", including least restraint, informed consent, and right to refuse treatment. These issues reveal the fundamental importance of patient autonomy as a value underlying the right to treatment. Despite this, the right to treatment is today at best empty—e.g., court reliance upon "professional judgment" in applying the right—and at worst distorted beyond recognition—e.g., appeals to the "right to treatment" as justifying *forced* treatment. Beyond minimally decent care in institutions and routine medical treatment, perfunctory constitutional recognition of the right to treatment has not translated into the right to *services*, nor into a right to services in the *community*.

We question the usefulness of *legal rights discourse* as a means of satisfying user "needs", when such discourse excludes the economic and political dimensions which define and shape rights and especially access to them. Power is a mediator of rights: it is not surprising that the right to treatment became an empty right and even turned against its "holders", given that users had little influence in its development and interpretation. Rights discourse has not succeeded in protecting the autonomy of the most vulnerable users. The attempt to create a right to treatment, a positive right involving a claim to receive something from the State, rather than a claim for a limit to the exercise of State power, has had the perverse effect of masking the more fundamental right to liberty, upon which the right to treatment finds its most convincing constitutional basis.

It is in the *right to liberty*, based on the widely-shared value of autonomy, where we find the State's obligation to provide the services that could improve the autonomy of involuntary patients and aid their transition to community living. An *enlightened* interpretation of the State's *parens patriae* responsibilities would not only respect but indeed be derived from this most fundamental right. The *parens patriae* role ensuing from the right to liberty aims to enhance the *personhood* of persons unable to participate fully in the polity and society due to reduced competence and capacities, or because of oppression, discrimination and poverty—all of which affect many psychiatric patients. Hence, the State is obliged to make an effort to make the right to treatment and other "rights" *meaningful* for those *least capable of exercising* their rights.

"Meeting the Needs of the Mentally Ill"

A Case Study of the "Right to Treatment" as Legal Rights Discourse in the U.S.A.

Introduction

This article examines the development of the "right to treatment" in the United States, and the changes in how this issue has been framed in recent years. We highlight the major themes and controversies, analyse the crucial concepts in terms of points of agreement and divergence, place the early and recent literature in historical context, and finally add our own thoughts bearing on some of the issues raised. Ethical and political considerations are emphasized, supplementing the far more common jurisprudential or clinical analyses.

The right to treatment was the first patient right to receive sustained attention in the literature dealing with mental health issues, and the first to be litigated. The issue achieved prominence in the literature as a reaction to litigation initiated by Birnbaum (1960, 1974) which, according to the first official reaction in 1967 of the American Psychiatric Association (APA), raised concern about inappropriate interference in the therapeutic context and professional managerial affairs (Council, 1967).

Prior to the first major successes in courts, while there were periodic scandals, films and critiques which raised the plight of persons in asylums as a moral issue, interest was short-lived and failed to focus on the need for systematic revisions of explicit duties, responsibilities, or rights. Reforms within the mental health system were difficult to achieve or sustain since they depended upon the charitable, paternalistic good intentions of non-patients (see Rothman, 1980).

A history of the elucidation of the right to treatment shows that other rights, including least restraint, informed consent, and right to refuse treatment, began to find related footholds, following deepening ethical and legal analysis of issues related to treatment of committed patients (Rachlin, 1988). We will argue that these issues, having been given life by the right to treatment, reveal the fundamental importance of patient autonomy as a value underlying the right to treatment.

An enlightened perspective of the State's parens patriae duties on behalf of vulnerable persons would not only view the "right to treatment" as consequential to the liberty interests of patients, rather than a paternalistic view of patient "needs", but would oblige the State to make an effort to make the right to treatment and other "rights" meaningful for those least capable of exercising their rights.

The Right to Treatment as an American Constitutional Issue

Psychiatrists — as practitioners rather than custodians — played a virtually non-existent role in the lives of public asylum patients during the 1960s. Psychiatric staffing of public mental hospitals was far below APA's own standards — themselves a compromise with what was thought feasible at the time (Birnbaum, 1960). The early right to treatment cases, far from representing litigants' disagreement with the nature of medical treatment, revealed not only lack of treatment, lack of treatment plans, and lack of physician notations in patient files, but also lack of even minimal care (Treffert, 1970).

In one typical early right to treatment case, the ratio of state hospital inmates to physicians was over 500 to one; this ratio within a hospital was worse than that enjoyed by residents of the area surrounding the institution (Birnbaum, 1969), and also less than that of New York's Sing Sing Prison. In the early 1970s several U.S. states provided only one unlicensed physician per 800-900 patients, and pellagra, a vitamin deficiency disease, was found in all Maryland state hospitals (Birnbaum, 1974). The key right to treatment case, Wyatt v. Stickney[28], revealed a patient staff ratio of 5,000 to one (Schwitzgebel, 1973).

Defendants in the early cases litigated by Birnbaum nevertheless asserted that patients were receiving "milieu therapy": at least under the conditions being examined, this amounted to a statement that the therapeutic process consisted of living in the four walls designated as asylum and controlled by custodians,

---

[28] Wyatt v. Stickney, 325 F. Supp. 781 (M.D. Ala. 1971).

whether or not a patient received treatment. Governments were institutionalizing vulnerable persons for a variety of reasons. However, whether patients in state mental hospitals were there because of senility, mental retardation, neurosis, delusions, or physical violence, they were subjected to very similar circumstances with respect to restraints, crowding, behaviour control, and drugging (see Albee, 1968; Birnbaum, 1969).

Medical codes of ethics have always been centered upon the obligations of individual physicians to individual patients; they could not ensure that there would indeed be a physician available for that patient, nor that asylums would be sufficiently funded to provide human living conditions. By the 1960s, conditions had deteriorated in the asylums to such a degree that it became obvious that the interests of asylum patients could not be assumed protected due to a practitioner-patient relationship which, in many cases, was virtually non-existent.

Morton Birnbaum's seminal 1960 article (the first dealing with a "right to treatment") noted that "although our society undoubtedly recognizes a moral right to treatment", "our law has not recognized this legal right" (p. 499). He viewed the warehousing of patients with virtually no care of any kind as a consequence of insufficient resources which society could, if it wished, rectify, and hence was fundamentally a <u>philosophical</u> rather than economic problem. Hence the irony that while mental patients were among the least vocal and the most ignored of all categories of patients in the health care system, their right to treatment became the first health care "bioethical" issue[29].

The literature on the right to treatment has always been almost exclusively American. This reflects the American tendency to frame ethical issues as rights to be litigated as constitutional matters in the courts. This debate has nevertheless greatly influenced the way in which governments, academics and the publics in

---

[29] The first right to treatment litigation predates the issue of informed consent in medical research, identified as the first major bioethics issue by Rothman (1991). Concerns over consent gradually developed in the 1960s pertaining to medical research subjects. Rothman dates the emergence of bioethics as a discipline, and the use of the term, from the 1970s. He noted that prior to the emergence of the "dual role" of physician as both care-giver and researcher, the interests of patients were assumed — too blithely in more recent years — to be protected by the mutual trust built up over years by a family and community for practitioners who were well-known to them.

other western countries conceived ethical issues in the mental health field. The rest of this part will therefore briefly summarize the framing of the ethical issues underlying the right to treatment as a constitutional "rights" issue in the United States.

Due Process: Fulfilling the Promise of Treatment

Early arguments characterized the right to treatment of involuntary[30] patients as a matter of due process under the U.S. Constitution: that an action taken to deprive an individual of his or her liberty must, as a minimal

---

[30] Ethical or legal claims to mental health services for voluntary patients revolve around substantially different issues: the tenor of that debate is rather what society or governments should do, in a context of scarcity, for reasons of efficiency or morality rather than what they are obligated to do according to the most fundamental (constitutionally protected) standards of fairness. To avoid confusion, claims predicated upon "rights" for voluntary patients are referred to in the literature as a "right to health care", a "right to community services", etc. See, for example, pp. 275-279 of Beauchamp and Childress (1989), and Pellegrino (1978).

Szasz (1969) saw the "right to treatment" as meaningless or coercive in the involuntary context, and as a claim in the voluntary context. He proposed instead a right to health which, more in keeping with the meaning of "right", "would obligate the State to prevent individuals from depriving each other of their health ... [and] would obligate the State to respect the health of the individual and to deprive him of that asset only in accordance with due process of law..." (p. 74). This suggestion seems well worth pursuing. However, in our view Szasz's argument with respect to the right to treatment fails to appreciate the usefulness of an analysis around right to treatment as predicated upon involuntary commitment: 1) if commitment were assumed to be justified — or even if it were an unjustifiable reality — does the State owe an obligation with respect to the person committed that it might not owe otherwise? 2) What does the State's obligation reveal with respect to the rationale for commitment? It is sometimes difficult to evaluate the premises of a proposition before we have delineated their logical consequences — at which time it might be easier to work back to the premises. This process can be helpful in revealing tautology or mistaken premises. For example, if the premise for commitment is assumed to be need for treatment, and we find that treatment is not provided, we might either follow our own logic and provide treatment, and then evaluate the results of this process, or realize that need for treatment was not, in fact, the premise.

requirement, bear a reasonable relationship to a rational legislative State purpose, and can be shown to be so in a process designed to ensure that liberties are not curtailed except insofar as necessary for that purpose. Mr. Justice Frankfurter wrote in 1950 that

> ... the Due Process Clause embodies a system of rights based on moral principles so deeply embedded in the traditions and feelings of our people as to be deemed fundamental to a civilized society as conceived by our whole history. Due process is that which comports with the deepest notions of what is fair and right and just. The more fundamental the beliefs are the less likely they are to be explicitly stated. But respect for them is of the very essence of the Due Process Clause ... (quoted by Birnbaum, 1960, p. 503).

If an individual is committed according to a criterion of "mentally ill and needing treatment", then the law clearly implies the "legislative promise" of treatment, which could be enforced as a statutory requirement (Schwitzgebel, 1973). This was the case with the first successful but limited right to treatment case, Rouse v. Cameron[31]. That decision, although based on statutory grounds, implied that a constitutional interest may exist. It did not require court determination of the nature or evaluation of the effectiveness of the treatment, but rather that a bona fide effort had been made, on the basis of the suitability of the treatment for the individual patient, and adequate in the light of present knowledge. Satisfaction of the court's requirements would take as little as a reasonably documented treatment plan, formulated in concert with recognized professionals and in accordance with professional standards, and for which implementation and monitoring can be demonstrated (Bazelon, 1969).

Stone (1975) characterized the right enunciated in 1971 by Wyatt v. Stickney, the first mental patient right to treatment case decided explicitly upon constitutional grounds, as "whether the hospital has made a reasonably convincing effort and whether the patient has received an adequate amount of a mode of therapy that some reputable segment of the profession deems appropriate"; hence the right would not imply the court choosing between therapies. The role of the court would therefore not be different than its role in reviewing bureaucratic agency action in other areas (p. 1128).

These major decisions aimed to respect the clinical independence of mental health professionals, while "it is for a court to decide whether the course of care in

---

[31] Rouse v. Cameron, 373 F. 2d 451 (DC Cir. 1966).

a particular case is legally sufficient to justify deprivation of the individual's liberty" (Halpern, 1969, p. 131). This quotation presages the degree to which the right to treatment will revolve much more around questions of patient liberty than around the nature and degree of treatment. Those decisions, and the importance to the courts of the role of "professional judgment", could be viewed as enhancing the role of the clinician, vis-à-vis the institution for whom the clinician worked. It would seem logical that the professional ethics and practice standards of psychiatrists, as well as their own corporate interests, would have incited the profession to lead the call for a right to treatment.

While some individual psychiatrists and the American Orthopsychiatric Association aided in developing the concept of adequate treatment, the APA and hospital psychiatrists firmly resisted the development of a right to treatment which, according to Halpern (Burris, 1969), "betrayed an acute desire to protect the psychiatrists' preserve against nonpsychiatric interlopers ..." (p. 128). Such an attitude would seem to indicate a lack of understanding as to the problems giving rise to litigation, and the nature of the remedies sought. The claims and court decisions would not have intervened in already established programs of psychiatric care, but rather require treatment where there had been none, or weigh treatment effectiveness against a statutory need for treatment criterion for involuntary confinement (Schwitzgebel, 1973).

A possible explanation for the opposition of most of the psychiatric profession to early developments in the right to treatment might lie in the role of psychiatrists as asylum administrators. As administrators in the public sphere, they may have felt obliged to not oppose the position of their employers; administrators or owners of private asylums may have feared spill-over of government or court regulation into the private sector.

However, this opposition may simply have been based upon ignorance of what the right to treatment would imply for psychiatric practice (as opposed to institutional and government policies). This was Stone's view; he regretted the passive role of psychiatry, since the memorandum of agreement ensuing from the landmark Wyatt v. Stickney case "does not reflect the self-interests of psychiatry or its special role among the mental health professions" (1975, p. 1127).

Stone's argument, which explicitly undertook to rebut the APA's "overstated" objections, recognized the constitutional right to treatment and demonstrated the need for it as a result of the failures of state administrations, concluding that "psychiatrists would do better to cooperate and support sensible litigation rather than become defensive and appear to be trying to justify substandard treatment (p. 1132).

Quid Pro Quo: Something "More" Owed by the State

The right to treatment first received the attention of the U.S. Supreme Court in 1975 in O'Connor v. Donaldson[32]. That decision seemed to base a right to treatment on the theory that persons confined under the parens patriae power should have the benefit of a quid pro quo: that removal of the rights of a person who is guilty of no crime is such a severe invasion of that person's integrity that the State owes some undefined "more" as a kind of compensation.

At minimum therefore, exercise of this power should not leave the patient in a worse condition than would otherwise occur if not committed. That very minimal responsibility of the State had been clearly abridged, however, in most of the early right to treatment cases, where patients suffered from malnutrition, violence (including from ward attendants), lack of activity, poor hygiene, routine debilitating physical restraints and drugging, etc.

There is no consensus whether the quid pro quo would or should apply to those confined as dangerous to others (Kapp, 1994). The justification for removing the liberties of a mentally ill person because of "dangerousness" to others, rather than because of guilt under criminal law, is that person's "non-responsibility" for the harmful acts. One cannot justify punishing a person lacking mens rea (a guilty mind) with poor treatment, unnecessary restraints and substandard living conditions.

Furthermore, since without treatment and reasonable care a non-criminal individual confined for dangerousness to others or incompetence to stand trial may well remain confined for longer than if that person had committed the feared crime(s), fairness would seem to imply those conditions and effective treatments reasonably required to enable reduction of that person's dangerousness and subsequent release (Schwitzgebel, 1973). Hence the logic of compensation for

---

[32] O'Connor v. Donaldson, 422 U.S. 563, 576(1975).

liberty deprivation supporting a quid pro quo under the parens patriae power would seem to apply equally to those mentally ill persons confined as dangerous to others.

The Supreme Court seemed to have backtracked on the quid pro quo theory, as such, in 1982 in Youngberg v. Romeo[33]. That decision took an important step by recognizing that patient liberty interests survived confinement, and that a duty might be imposed upon a state by a federal court on constitutional grounds. The court greatly qualified these principles, however, by stating that patient liberty interests could be balanced by State interests, and in doing so courts must show deference to the judgment exercised by a qualified professional (unless the judgment substantially departs from professional standards). The decision recognized the principle of the right to treatment, but required only minimally adequate or reasonable treatment or training to ensure safety and freedom from undue restraint (Seicshnaydre, 1992).

Youngberg v. Romeo therefore does not explicitly provide a clear constitutional rationale as to why the State should do something "more" for involuntarily committed mental patients. However, Stefan (1992) believes that the rationale of Youngberg implies "treatment" favouring autonomy: "Under Youngberg's rationale, the constitutional predicate for the right to treatment is to protect or increase liberty, including liberty from the total restraint of institutionalization" (pp. 688-689).

While such a rationale might be implied, both that writer and Perlin, Gould and Dorfman (1995) have viewed the way in which the court has interpreted the "professional judgment" standard as effectively rendering treatment as well as other patient rights unenforceable. These latter writers applied the new "therapeutic jurisprudence" analysis to Youngberg, concluding that the decision ignores or distorts published research and is "profoundly antitherapeutic" (p. 105).

State Interests and Paternalistic Duties

Unless there is a statutory basis for a claim, it would appear to rest with a very paternalistic interpretation of the State's parens patriae obligations, which

---

[33] Youngberg v. Romeo, 457 U.S. 307 (1982).

themselves may be balanced with "State interests". Furthermore, since the ways in which most commitment laws are written and applied does not require specifying in individual cases whether the person committed is considered a danger to self (parens patriae) or a danger to others (police power), or both, it is difficult or impossible to sort out the State's obligations and justifiable interests (see Kittrie, 1971, pp. 394-400). The Youngberg judgment can therefore result in the patient's interests being "balanced" with an unjustifiable State interest, e.g., protecting the public, when the individual patient is not in fact dangerous to others.

The constitutional basis for a right to treatment according to the due process argument is convincing in the limited case where a patient is committed under a statute that justifies commitment on the basis of the patient's need for treatment. However, the tendency in western jurisdictions during the late 1970s and 1980s to make commitment criteria seem more restrictive by adding dangerousness criteria (Aviram and Smoyak, 1994) has had the perverse effect of compromising the statutory basis for a right to treatment. The due process argument alone is difficult to sustain on a constitutional basis when the state rationale for confinement is not necessarily tied to need for treatment.

Adding to the soup a quid pro quo argument has not helped. While that idea may hold a certain intuitive appeal, its constitutional as well as ethical basis remains murky, and its implications for treatment rights even more so. As noted before, it is not clear that the U.S. Supreme Court still holds to that theory. Stefan (1992) suggested that exchanging treatment for liberty may not be a fair trade. One might also argue that the State owes compensation only if culpable and that fairness lies not in the outcome of confinement length, but rather in ensuring that non-criminal patients receive treatment at least equal to those of criminals.

Rights to Community Services

The litigation approach — and more specifically with respect to treatment and care in institutions — was not only controversial among clinicians and asylum administrators, but also among reformers. Many civil libertarians and civil rights organizations felt that attempts to improve care in institutions implicitly endorsed asylums as opposed to community care (Birnbaum, 1974).

Interestingly, those holding this view later found strange bedfellows among psychiatrists such as H. Richard Lamb and E. Fuller Torrey who are well-known for their opposition to expanding civil rights for patients and to the speed at which institutions ejected patients into the community. An APA task force (chaired by Lamb) issued a statement to the effect that "legal advocacy on behalf of institutionalized mental patients 'neglected [the patients'] right to high-quality comprehensive outpatient care'" (quoted in Perlin, 1994, p. 202).

However, the history of the right to treatment shows that the reverse is true. "Virtually every early case explicitly sought an expansive provision of community services including after care, transitional services, and suitable community-living arrangements" (Perlin, 1994, p. 201). In 1976 Dix wrote that despite recent pessimistic criticism, the right to treatment decisions were flexible enough to open the way to procurement of community care resources. Furthermore, lawyers who litigated those cases (e.g., Birnbaum, 1969) felt that improved standards in asylums — and hence higher costs — would provide incentives to governments to reduce asylum confinement for patients for whom community care would be more cost-effective — hopefully freeing funds for community care (Kapp, 1994).

This theory was tragically wrong, as is well known today. While higher institutional standards did indeed cause further deinstitutionalization (Stone, 1975), and forced policy planners to redirect their attention to community care, the increasingly restrictive fiscal environment greatly reduced hoped-for funding for the new programs. In this context, "any public consensus that had supported expanded civil rights and discretionary programs for mentally disabled persons was sharply eroded", along with a sharp reduction of funds for social-service programs and increasingly negative attitudes toward people with mental disability (Perlin, 1994, p. 195).

Furthermore, common with health care in general, the system's inertia has made it very difficult to transfer resources from hospitals and asylums to service or prevention oriented programs within the community (Miller, 1991; Aviram et al., 1995). Some of this inertia is attributable to a mindset noted by Albee in 1968 which insists that the key element of help for distressed persons is psychiatric services: yet there can never be enough psychiatrists, and few psychiatrists are willing to practice in those community settings where the incidence of distress is highest.

Despite some early lower court successes in the United States, recognition of asylum inmates' right to treatment has had little or no constitutional impact with respect to community services. Legal decrees extending the right to treatment into the community tend to have been based on interpretations of statute law[34]. The Supreme Court has shown what several critics consider to be such excessive deference to professionals that courts are prevented, in effect, from judging cases on the merits of patient liberty interests (Perlin, 1990; Stefan, 1992).

Consequently, according to Perlin (1994), litigators, legislators and scholars gave up on the issue of the right to community services. While innovative litigation with respect to the right to treatment has "slowed to a trickle", this "stands in stark contrast to such other areas of constitutionally based mental-disability law as the right to refuse treatment, in which caselaw developments continue to increase exponentially" (p. 200).

On the other hand, litigants for patients have been successful in some cases where the professional judgment exercised was that the institutionalized patient should be released to community care settings which did not exist. A proposed treatment and care plan that excludes services simply because they are not available has not been considered acceptable professional judgment (Seicshnaydre, 1992).

This is highly significant, because in the past much reform has been stymied by the tendency of medical practitioners to work within the system as it is rather than attempting to change the system. For example, New Jersey's new screening services, established to aid transition of users to less restrictive environments in the community, have instead tended to encourage hospitalization too readily, due to limitations in available community resources. Such services, according to Aviram (1993), threaten to become "gate openers" to hospital commitment rather than "gate keepers".

With respect to empowering and normalizing care and services in the continuum from asylum walls to community living, the most promising constitutional directions today were originally formulated in the context of right to

---

[34] An interesting example is provided by Arnold v. Arizona of 1989, which ordered the creation of a "comprehensive, community-based system of care for 4,500 indigent and chronically mentally ill individuals" (Petrila, 1992).

treatment litigation but have achieved conceptual and constitutional independence from the "right to treatment": "least restraint" and "equality rights".

Least Restraint. While the Supreme Court has never accepted the principle of "least restraint", the language with respect to "undue restraint" is a step in that direction (see Appelbaum, 1987). This implies that caregivers must be prepared to justify specific restraints on liberty, and secondly, that they are reasonable. Hence, substantial restraints for mere administrative convenience might be unconstitutional.

While the right to treatment has not extended to voluntary community services per se, it seems that requirements for "reasonable" care without "undue" restraint might be used to create community services necessary to enable involuntary institutionalized patients to return to the community[35]. A government obligated to provide a community care infrastructure for former involuntary patients will also have some incentive to extend services, as a preventive measure, to others who otherwise might end up in institutions.

Equality Rights. Equality rights statutes, such as the Americans With Disabilities Act (ADA), may enhance the constitutional arguments of advocates of least restraint and community services, on the grounds of equal treatment requiring reasonable accommodation (Seicshnaydre, 1992). Perlin (1994) reports an analysis by Cook which indicates that segregated settings for disabled persons are unacceptable under the Act. Of particular promise in the United States is the phrase in the ADA which stated that the Act provided a means to enforce the equal rights provision of the Constitution; "This means that any violation of the ADA must be read in the same light as a violation of the Equal Protection Clause of the Constitution" (p. 203).

The ways in which the ADA will be used by rights advocates in the United States will be of interest to all countries, since statutory definitions can be used to

---

[35] An important article by Chambers developed constitutional arguments for the "least restrictive alternative" and "somewhat tender constitutional arguments" (1972, p. 1183) requiring community services with respect to those who would otherwise risk being committed, de facto or otherwise.

clarify applications of constitutional provisions. The equal rights route might someday, in a more favourable political and economic environment, open the door to services to voluntary, as well as involuntary, patients in the community in order to prevent the segregation from society implied by institutionalization.

The Mitigated Success of Litigation

Stone (1975) characterized the problems giving rise to right to treatment litigation as a matter of access to health care:

> In the end the real solution to the problems addressed by the right to treatment cannot come from complicated judicial discourse about civil rights and civil liberties. It must come in the form of a system of national health insurance that includes adequate mental health coverage for inpatient as well as outpatient treatment and for chronic as well as acute mental illness (p. 1133).

On one level Stone is undoubtedly correct: in the United States as elsewhere there is less than perfect "parity" between coverage of mental health services and other health services; for many years U.S. Medicaid excluded state mental hospital residents under the age of 65; and the U.S. health system leaves many without insurance coverage. However, the circumstances of the severely mentally ill were so degraded with respect to others in society that there were more fundamental factors that contributed to their disfavoured health insurance status.

Wald and Friedman (1978), in an excellent but under-appreciated early history of the development of patient rights in the United States, demonstrate how the "tide of litigation", while according to some critics an unjustifiable invasion of the prerogative of governments to set policies and allocate resources, amounted to the only opportunity for patients of mental institutions to advance their demands:

> The mentally ill have not traditionally been very successful in displaying clout with politicians or legislatures; their advances have been grounded in basic principles of fairness and equity, traditionally the province of the courts. If the power of courts to influence the mental health systems is cut back they will almost surely suffer (pp. 141-142).

Despite the several lower court right to treatment successes, however, Stone noted the highest court's blind spot with respect to asylum patients, and its implications for the pursuance of patient rights:

> It is interesting that not one [of the ground-breaking cases] arose in the context of the more numerous and familiar cases of civil commitment of the mentally ill. All of the cases involved men who ... had been originally charged with crimes; they therefore had extensive access to legal counsel. This perhaps is illustrative of the lawyers' contention that without the right to counsel all other rights are bootless. (1975, p. 1127)

Hence the litigation has evinced a perfunctory constitutional nod toward the "right" of committed persons to be maintained at some minimal level of health and well-being, subject to State interests, but if there is also an obligation to provide the kind of treatments and services that would enable the patient to rejoin the community, it has not been made clear by the U.S. Supreme Court. The arguments, along with the claims for equal protection before the law, were finally, after years of litigation, sufficient to claim for asylum patients living standards and treatment rights at least equal to those of criminals — and laid the groundwork for further arguments with respect to least restraint and due process issues related to civil commitment.

Given the frustrations experienced by lawyers who had litigated right to treatment cases, therefore, several supported a legislative approach. However, it is easy to see today that many of the mechanisms subsequently put in place by legislation to enforce patient rights have been ineffective, and may indeed be less effective than was the constitutional litigation of the 1970s. Even when formally recognized, patient rights remain very difficult to enforce; sanist attitudes can afflict the courts and patients' lawyers (Perlin, Gould and Dorfman, 1995). This has resulted not only in rushed patient hearings and minimal due process protections (Cohen et al., 1997), but also in the routine bypassing of those procedural protections clearly required by statute and court decisions (Wald and Friedman, 1978).

However, Perlin (1994) and Kaufman (1979) believed that the litigation was highly successful with respect to improving institutional standards, as most institutions finally complied with most of the care standards sought by the litigants. Apparently the litigation created expectations which are now widely reflected in both statutes and institutional practice, although it must be recognized that extensive reforms of the system had already begun as a result of a wide variety of social, ethical, political and economic changes sweeping society, as well as changes in the practice of psychiatry and other health and social services (Kunjukrishnan and Varan, 1989; Rappeport, 1987).

Indeed, the interest of some lawyers litigating the right to treatment, the arguments they drew upon, and the growing receptiveness of courts and mental health professionals to these arguments, were greatly influenced by these changes (McCubbin, 1994). It may well be that litigation was but a symptom of a process that would, litigation or not, have drastically improved the respect for patient rights and the quality of care to institutionalized patients (see Leaf, 1977).

The right to treatment seems today more significant for laying the groundwork for recognition of mental patients' inherent liberty interests than for improving access to quality care for psychologically distressed persons. By breaking down the justifications for deprivation of liberty in order to establish a right to treatment, a Pandora's box of ethical and constitutional concerns was opened. In the development of the right to treatment, therefore, we find that the nature of the existing power relations led patients and their advocates to advance their claims through the court system by using a rights discourse, which has provided unquestionable though limited success.

The Right to Treatment as Legal Rights Discourse

The "Right" to Forced Treatment

Advocates for commitment criteria based only on "need for treatment" have increasingly appealed to the "right to treatment" as a justification for forcing treatment[36]. This phraseology seeks its justification in the idea that a patient may

---

[36] Witness, for example, the oxymoron contained in the article title "The dilemma of civil rights versus the right to treatment ..." (Isohanni et al., 1991). They wrote:

> The doctors have given such priority to the need-for-care argument that, in some cases, they have admitted involuntarily for assessment patients who have been in obvious need for care, but did not meet the diagnostic criteria in the law. ... The doctors' view seems to have been shared by the patients and relatives, because no one made formal complaints about these admissions. ... The patients' right to the best available treatment is an ethical aim that finds support from doctors but simultaneously conflicts with the legal aspects of civil rights. This discrepancy seems to be leading doctors, in some cases, to conscientiously object in interpreting the law (pp. 260-261).

be incompetent to refuse a proposed treatment: that the refusal is a symptom of the illness which needs to be treated. Therefore, pursuant to its parens patriae role, the State has an obligation to act in the "best interests" of the patient by forcing treatment. Interestingly, the issues of forced rights to treatment and the extension into the community of services to which patients have a right converge with respect to patients subjected to the recent development of outpatient treatment orders.

We take issue with characterization of forced treatment as a matter of the patient's "right to treatment". Firstly, forced treatment does not require for its justification a right to that treatment; secondly, such justification becomes problematic given that it would seem to conflict with other human rights such as privacy and autonomy (Hermann, 1990; see also: Derdeyn, 1977; Froese, 1991).

If forced treatment could ever be justified on the grounds of the good of the patient, there is a developing consensus in the ethics and jurisprudence literature that it would have to be predicated upon the patient's incompetence, the long-run net benefits for that patient of various alternatives, and an estimate of what, given the patient's basic character and preferences, that patient would decide if temporarily competent (McCubbin and Weisstub, 1998). By definition, a "right" is a very powerful claim which cannot be compromised except by other rights. Hence language which characterizes forced treatment as responding to a "right" held by the patient must render as secondary the value of patient autonomy, trivializing the patient's attitude to the treatment.

Furthermore, the mere fact that a patient objects to a treatment reduces the net benefits of imposing it (McCubbin, 1997). Even if an incompetent patient's evaluation of the proposed treatment is completely fanciful, the act of forcing it is hurtful to the patient and may have a disempowering effect. This provides a

---

Szasz (1969) had in effect predicted this result years before: he opposed the "right to treatment" because, among other reasons, the issue arose with respect to involuntary patients and that therefore the required treatment would have to be forced. "[W]hose right to treatment do the advocates of this concept wish to guarantee — that of the parent to commit his rebellious son as mentally ill or that of the child to defy his parents without being subjected to quasi-medical penalties?" (p. 71). Nevertheless, it is also clear that the right to treatment litigation was motivated by and raised questions of patient liberties with respect to commitment and treatment refusal (see McGough and Carmichael, 1977; Wald and Friedman, 1978).

complex problem which seems to be ignored by the simplistic formulations which are often heard which more often seem to equate the patient's "right" to treatment not with what the patient wants but with what the patient "needs" (see Charest, 1994; Lamb and Mills, 1986; Weitzel, 1977). Such writers tend to perceive fundamental conflicts between the right to treatment and the right to refuse treatment (Kleinman, 1986). We find those arguments disingenuous since they logically imply that a patient demands and refuses the same treatment.

Miller (1991) suggests that conflicts over moral obligations of patients and clinicians have been confused with conflicts over legal rights. In fact, in this case the "conflicts" arise when someone attributes a demand to a patient for a treatment which the patient refuses — or, as might arise in a case of euthanasia, where a substituted decision maker attributes a refusal of life-sustaining treatment to a patient who had demanded it (Gunn, 1991). If a patient is incompetent to refuse a treatment she or he is also incompetent to demand a treatment.

Hence the "conflict" resolves not around the rights but rather issues of informed consent and coercion (see Moorhouse and Weisstub, 1996), the determination of incompetence[37], and substituted decision making (McCubbin and Weisstub, 1998). It is in those issues that we should seek the roots of a supposed conflict between the right to treatment and the right to refuse.

It is much more common today to hear the "right to treatment" expressed in the sense of the above paragraphs than in the original sense. The remainder of this paper will explore the reasons for that, by examining what a "rights discourse" means for mental patients in contemporary cultural, economic, legal, and political contexts. We will suggest that even a "successful" — popular — rights discourse is liable to have little more than symbolic value for the "holders" of those rights, when the holders have little influence in problem definitions that give rise to their

---

[37] Beck and Parry (1992) suggest that forced hospitalization or treatment should not be based on mental illness, "need for treatment", and/or dangerousness, but rather incompetence, where the person has a severe mental disorder and cannot live safely in freedom. Obviously, the meaning and assessment of "incompetence" is a thorny issue (Winick, 1996). Clearly the issue of competence should be decided independently of the patient's preferred treatment decision, or the process becomes tautological (considered incompetent due to "lack of insight", evidenced by disagreement with physician's diagnosis or treatment plan).

expression, little influence in the way claims are framed (as rights or otherwise), and, especially, little influence in how they are to be enforced.

Mental Patients as a Disempowered Group: Ethics, Rights and Realities

The fact that virtually the entire discussion of the "right to treatment" throughout the history of this literature revolved around interpretations of constitutional legal rights, with virtually no attention to the issue as an ethical issue, per se, is understandable when we realize that marginalized groups in general, but particularly severely distressed individuals, enjoy relatively little influence upon the general public or the executive and legislative branches of governments which are politically responsible to them.[38] This was recognized in the ADA by the U.S. Congress, describing disabled persons as a "discrete and insular minority ... subjected to a history of purposeful unequal treatment, and relegated to a position of political powerlessness" (quotation from Perlin, 1994, p. 203).

Isolation. Almost half a century ago Weihofen and Guttmacher provided a thought-provoking explanation for the isolation of those labeled mentally ill from society:

> Most people defensively feel that insanity can never come to them. An old term for insanity, "alienation", clearly portrays this. ... What may happen to me is more important than what is happening to others (1952, quoted by Birnbaum, 1960, p. 502).

This suggests that it may be difficult for many to perceive the "rights" of the mentally ill as a matter of "justice" as conceptualized by Rawls (1971), who studied

---

[38] For a detailed discussion of how those labeled mentally ill are politically handicapped see McCubbin and Cohen, 1996. In a case study of the process of changing New Jersey's commitment laws, Aviram and Weyer (1996) identified the major actors, none of whom were users, and how the "arguments presented in terms of what was or should be in the best interests of clients actually reflected organizational and professional concerns regarding power and turf" (pp. 798-799). In adapting to environmental changes, these actors deflected the "reforms" from an initially strong libertarian orientation to a social service approach.

the problem of what rules would be favoured by individuals who are ignorant of their particular interests, tastes and resources, and of their situation within society. If in that "original position" of ignorance about what we will become we cannot envisage the possibility of becoming mentally ill, we may neglect to seek a social contract that "insures" against this eventuality. This helps explain Aviram's observation that "the issue that concerns society and draws public attention and moves governments into action, is not treatment versus liberty, but social control versus deviant behavior" (1990, p. 175).

This isolation, whether due to fear, hatred, need for treatment, or control of social deviance, is very usefully understood as an epistemological problem. What happens to our ability to "know" something by objectifying, distancing and isolating it? Do we change the nature of the phenomenon we wish to observe by so doing? In the case of persons we label mentally ill, what does institutionalization imply within this problematic? Is such a social institution liable to reproduce itself; a dynamic which becomes reinforcing as we lose the capacity to observe mental illness within its "natural" environment? Does this imply a growing socially <u>constructed</u> concept of mental illness in which the "real" phenomena become increasingly heavily overlaid with attributed phenomena?[39]

Consumer Rights. On the other hand, since as the years pass more and more people are considered by mental health epidemiological research to be mentally ill, and more and more people are labeled by others or themselves as being

---

[39] Perhaps the best discussion of this as an epistemological problem is not Foucault's obvious classic but rather a little known work (1962). See also: Jennings (1994) and Larner (1995).

Referring to Adorno and Horkheimer, Farber (1995) characterized the problem in terms of the dialectic of enlightenment:

> As such, it is confronted with an insoluble dilemma. On the one hand, it desires to annul the Other whose otherness poses the danger of disorder. On the other, it requires the existence of the Other. The ideology of service sets up a dichotomy between servicer and client which can never be overcome; existential security is sought through control and mastery of the Other, not through the recognition of otherness. Clients must be sustained in their otherness while at the same time being compelled to constantly re-enact the rituals of submission. They cannot be permitted to transcend or legitimate their otherness by acquiring the rights and powers of citizens, by becoming full subjects. In that case there would be no Other to manage control (pp. 189-190).

mentally ill, dealing with psychological problems, or suffering "chemical imbalances" in their brains for which they are taking psychoactive medications, Farber suggests that "the advance of the therapeutic state may carry within it the seeds of its own demise. When every citizen is at risk of being transformed into a client, the foundation of the therapeutic state becomes shaky" (1995, p. 192). While Farber's point is well-taken, we should recognize the growing split, in terms of interest representation in advocacy movements, between those in the most severe distress and those "consumers" having relatively mild problems (Everett, 1994; Mechanic, 1994).

It is an open question as to whether rights won by "consumers" will be of significance for those with the most severe problems or subjected to the most stigma and constraint by law, social attitudes, and clinical practice. The "consumers" might be regarded as the vanguard of a movement benefiting all patients, or they might coopt other patients' voices in a mental health system which becomes bifurcated: one for the "consumers" who act as marketplace buyers of services, another for coerced patients — some of whom may not wish to "consume" the forced services. It is this consideration which led us in this paper to refer to the involuntary clients of the mental health system as "patients" rather than as "consumers" or "users".

Rights as Rules. It is not essential that every widely acknowledged moral "right" or value become the object of constitutional protection. Other enforcement mechanisms may be available: the right of a spouse to be treated with respect by the other spouse can be, to some degree, enforced by the civil laws governing contract, divorce, and (perhaps especially) alimony. The right of a boy not to be spat upon by another boy has been traditionally enforced by "playground justice" (at least until the over-growing judicialisation of children's inappropriate behaviour). Within a small community the disapproval of neighbours, relatives and friends can help to control irritating swearing, smoking or drinking (or, on the other hand, tolerate it by an implicit pact of mutual non-interference). Unfortunately, in a large anonymous society, social reciprocity is less effective as a means of social control, so we have increasingly relied upon our governments to establish and enforce new rules.

Such rules may or may not succeed in achieving outcomes which are in some sense ethical or socially optimal, but every rule also creates perverse outcomes which give rise to ethical problems. In the context of mental health care, there can be no doubt that specific applications of patient "rights" can result in some patients being worse off in some sense. Analysis of the problem should take place at both micro and macro levels: Can the rule be refined to account better for the variety of circumstances in which it would be applied? Upon what does the rule derive its justifications? Is there a more general rule which should replace it? Is the rule, or the necessity for it, contingent upon certain assumptions or conditions which might be wrong or alterable?

Conflict and Alliance. Inevitably the application of a protected right to specific types of situations is a matter of interpretation requiring empirical analysis, and in the real world different rights are seen to conflict[40]. The process of mounting arguments which develop the facts viewed as pertinent and justifying the primacy of the claimed right over others requires resources of power — including money. For most mental patients, as a group greatly disfavoured economically, courts are inaccessible since most countries provide little or no access to legal help; a problem compounded by the social handicaps borne by many institutionalized mental patients.

This points to a paradox with respect to individually based "rights" in liberal capitalist countries, when those who "enjoy" those rights are members of groups who, in such societies, are systematically oppressed, detested, ignored, or otherwise marginalized to the edges of the mainstream. This paradox is illustrated by the irony that many of the patients litigating for the right to treatment, according to case descriptions found in Burris (1969), seemed less to require "treatment" than community services, and probably should not have been institutionalized in any event (see Albee, 1968; Twerski, 1971), while those who most needed treatment and psychosocial services in order to be able to live autonomously were also those least likely to succeed in marshaling the resources and cooperation necessary to litigate.

---

[40] This is discussed in Hayes, Boerma and Ovadia (1993-94), and in Schopp (1993).

Indeed, it is interesting to note that the most influential right to treatment case, Wyatt v. Stickney, was brought by laid-off asylum employees (Schwitzgebel, 1973). In polities which provide benefits passively through the balancing of competing rights and claims, the rights that are given substance for powerless groups are likely to be contingent upon the convergence of interests with those who have power. To the extent that meanings and operationalizations of rights are controlled by those powerful groups, they may be symbolic, or even perverse, as the mutation of the right to treatment into the right to forced treatment well demonstrates.

To the extent that the interests of the powerful and powerless may again diverge, even "substantive" rights will be temporary, unless there is change in those fundamental societal structures that create systematic, reproducing powerlessness. While these observations are obvious with respect to some oppressed groups — e.g. black persons in the United States — the paternalistic discourse of "caring for the mentally ill" or "meeting their needs" has allowed us to diminish the personhood of those we have so labeled. Only persons are allowed to exercise on their own behalf the "rights" that have supposedly been attributed to them.

Rights in the Community Context. Indicative of a growing "communitarian" critique of rights discourse, Seicshnaydre suggested that "rights analysis emphasizes individuality and autonomy and fails to recognize social relations and responsibilities between groups" (1992, p. 1997). Such recognition would imply the realization that not only do individuals and groups have differences with respect to needs, but that they also have differences in terms of capacities — the most crucial being their capacity to seek and enjoy that which is due them by right. We share, therefore, Seicshnaydre's conclusion that while a rights approach cannot be blamed for the slow progress of the establishment of community care, relations involving power and responsibility also have to be evaluated (see Cohen and McCubbin, 1990).

A hierarchical approach to rights might provide a useful way to address these relations (e.g., Schultz, 1996). Some rights are preconditions for the ability to enforce or enjoy others: specific rights for specific categories of persons remain empty, unsought or coopted, when those persons are unable to exercise their right to democratic civic participation — because of the power of others and because of their own mental or physical disabilities. The interpretation of facts is always coloured by the values of those interpreting them, and the sorting out of conflicting rights can appeal to nothing other than the values of those doing the sorting.

These values, consciously or unconsciously, are built upon some combination of self-centered "desires" (hedonistic, selfish, utilitarian, etc.) as well as upon what Fromm (1947) refers to as the "humanistic conscience". We might associate such a conscience with the ethical good, whether deontological or teleological. Contemporary attempts to naturalistically derive the ethical good have tried to replace mystical obligation with a kind of "higher level" utilitarianism which recognizes the social nature of happiness (see Bertalanffy, 1981; Etzioni, 1988; Frank, 1985; Maslow, 1970). Such formulations have included a sense of justice or efficiency which requires abstraction beyond identification as an individual to identification as community (e.g., Hegel, 1952; Kant, 1958; Loewy, 1993), an abstraction which may not be as different as some communitarianists have suggested from Rawls' original position (1971).

Power as a Mediator of Rights. Whether legal rights grow from or are justified as communal moral values, or whether they can be seen as the result of a rational or negotiated social contract (Locke, 1963; Rousseau, 1992), it is clear that a rights discourse alone will not resolve most real-world situations. Individual and group values enter into the equation not only as sources of legal rights but also as mediators of those rights (Bell, 1996). Power then becomes important, not only for advancing individual desires but also for disseminating or enforcing beliefs respecting ethical values, whether or not they underlie legally protected rights. Furthermore, even if for everyone legal rights were to be synonymous with moral rights, there would still be the question of what weights different groups and individuals place upon them when they conflict (see Ubel and Loewenstein, 1996).

Therefore, given the ever present importance of the role of power, and the fact that values are built upon both the more narrowly circumscribed desires and interests as well as more "humanistic" concerns — which we will consider here to be the subject-matter of "ethics" — implies a two-pronged concern for the application of ethics. That is, that ethical concerns enter not only with respect to the relative merits of the competing values, but also with respect to the nature of the process that weighs those competing values. Ethics is concerned not only with "good" results but also with "good" means (McCubbin and Cohen, 1997b) [indeed, many would assert that good or bad results, in themselves, have nothing to do with ethical value].

U.S. Supreme Court decisions have firmly and explicitly deferred to "professional judgment" in recent decisions regarding competency and refusal of treatment (Perlin, 1990)[41]; hence U.S. courts have been most receptive to right to treatment litigation on behalf of institutionalized mental patients when professionals themselves have recommended treatments or services that the state or institutions have not implemented (Seicshnaydre, 1992). Some writers have suggested that some early support by the mental health professions, particularly psychiatrists, for institutional reform, deinstitutionalization, and the development of a "community care" model facilitated if not enabled the changes which resulted, and that that support was likely based at least in part upon professional self-interest — e.g. more funding, better working conditions, higher status associated with a medical rather than custodial role (Brown, 1985; McCubbin, 1994).

Given the highly unequal nature of patient versus professional power, and the carefully phrased way in which the Supreme Court has interpreted and applied patient rights, one might question the degree to which patient rights "victories" are instead shaky partnerships which fall short of providing fundamental constitutional protections for the patients. Does the espoused "right" bear an unambiguous relationship to fundamental widely shared ethical values, or has its application resulted from a conjunction of specific circumstances contingent to a distribution of power by which the basic interests of patients may remain largely unprotected unless they coincide with the interests of others?

---

41. Stefan (1992) demonstrates the implication of the "presumptive validity" given to professional judgment by citing Justice Blackmun's opinion from a prison case:

> Such a requirement boils down to a command that when a court is confronted with a charge of administrative bad faith, it must evaluate the charge by assuming administrative good faith .... I regard it as improper to make the plaintiff prove his case twice by requiring a court to defer to administrators' putative professional judgment. .... I am concerned about the Court's apparent willingness to substitute the rhetoric of judicial deference for meaningful scrutiny of constitutional claims ... (pp. 692, 716).

In other words, to the extent that patients succeeded in achieving a right to treatment, was this a "good" result that luckily arose despite a deontologically "bad" process? To what extent does recognition of the right become moot if power relations prevent its enjoyment in the sense intended, or if it becomes empty because the system adjusts to it by altering the conditions under which it arises?

Identity Politics. There has been such an explosion of "rights" in recent years that there is a growing danger that the more politically powerful actors will be able to achieve dominance for their favoured rights, however trivial, over the claims of disfavoured groups based on values which are more fundamental. In this era of "identity politics" there is a risk of falling into relativism, where so many claims are justified by the act of having been put forward, that there is no longer room for societal reflection over the basic ethical values that should balance them (see: Feher, 1996; Glendon, 1991; Rosenau, 1994). Hence social choice is reduced to a game in which moral discourse becomes a cynical exercise known to be pretense by all players, since the winners are accepted to have the most resources and better strategies.

Conclusion

The Right to Liberty and Parens Patriae

As has been demonstrated in our discussion of the "right" to forced treatment above, rights discourse has not succeeded in protecting the autonomy of the most vulnerable users of the mental health system. We suggest that the attempt to create a right to treatment, a positive right involving a claim to receive something from the State, rather than a claim for a limit to the exercise of State power, has had the unfortunate effect of masking the more fundamental right to liberty, upon which the right to treatment finds its most convincing constitutional basis.

It is in the right to liberty, based on the widely-shared value of autonomy, where we find the obligation of the State to provide the care and treatment that could improve the autonomy of involuntary patients and aid their transition to normalization and community living. An enlightened interpretation of the responsibilities of the State under the parens patriae power would not only respect but indeed be derived from this most fundamental right. The parens patriae role ensuing from the right to liberty aims to enhance the personhood of those members of our society who are least capable of asserting it themselves.

A person who has lost liberty, for whatever reason, becomes dependent on State power, and therefore the State then must exercise a parens patriae role insofar as it doesn't conflict with another justifiable role (such as, as some would have it, punishment of a criminal). As with our contemporary conception of a good parent, the State role with respect to the dependent person is not only sustenance, and preventing deterioration, but also to favour development of that person toward autonomy.

This perspective is reflected in the U.N. General Assembly resolution, adopted in 1991: "Principles for the Protection of Persons with Mental Illness and for the Improvement of Mental Health Care"[42]. The principles seem to go well beyond that which has been explicitly required by U.S. constitutional law, calling for the establishment of community-based services and treatment in the least-restrictive environment, because "treatment shall be directed towards preserving and enhancing personal autonomy" (Rosenthal and Rubenstein, 1993, p. 261; emphasis added).

This implies not only treatment, but also minimal restraint, augmentation of the patient's decision making capacity and role in decisions affecting him or her, the provision of an environment — e.g. a community setting — that is the most fertile for the development of autonomy, and the provision of non-medical services that would foster competence, well-being, and growth toward independence.

Therefore, the "right to treatment" issue today is not a question that can be left to lawyers and judges. Social scientists, philosophers, and clinicians, as well as patients and ex-patients and the general public, have an important role to play in further clarifying the obligations of society to vulnerable persons — particularly when the same society, acting through its legislative, executive and judiciary removes or permits the removal of fundamental liberties.

---

[42] The World Health Organization has accordingly emphasized "consumer" empowerment and the role of consumers in psychosocial rehabilitation (Bertolote and Sartorius, 1996).

Toward the New Millennium: The Ethical Duties of the Helping Professions

In the early days of litigation for the right to treatment, when the demands of the litigants were highly reasonable given the scandalous conditions in the asylums, the APA and many psychiatrists reacted with defensiveness rather than an openness to face ethical problems going to the heart of their professional codes of ethics. This resulted in an explosion of legal and academic analysis of the mental health system, accompanied by a cascade of new laws and standards regulating the practice environment of mental health professionals. The history of the right to treatment shows not only how patients can be coopted while achieving this right, but also how asserting one right can facilitate the development of others.

Today, the mental health professions would do well to cooperatively seek common ground with patient rights advocates, taking care to avoid an interpretation of the Hippocratic Oath which distorts it into a mandate to control patients and health policy. This implies, according to Stambolović, replacing the focus on dominating the object with a focus on the development of relationships, since

> ... by protecting the human rights of the other, one is protecting one's own rights. .... Human rights cannot be reduced to the ethical intervention of scientists and/or humanists, neither can health care be reduced to technical procedures under the jurisdiction of experts (1996, p. 301).

It is not only an ethical obligation for the mental health professionals to listen respectfully to what users and their advocates have to say; it is in their long-term self-interest. Just causes have a way of eventually making themselves felt; their implantation can be sudden and radical if they had been frustrated for a long time. The detailed orders issued by Judge Johnson in the Wyatt v. Stickney case are a demonstration of this. According to Petrila:

> For the first time in the mental disability field, leaders are emerging from those most affected by decisions traditionally made by others. This is resulting in challenges to traditional notions about services, about legal principles, and about control of decision making in critical areas of people's lives. It is too early to predict the ultimate impact of these movements ... (1992, p. 102)

Practitioners now have an opportunity, indeed, an obligation, to take a much larger psychosocial perspective to patient needs, a point made by Wolf in calling for a "new ethics" in which the "physician should think systemically" (1994, p. 34)[43]. If courts are reluctant to face the issues thereby raised they have provided grist for a rapidly multiplying and diversifying literature by clinical ethicists, philosopher ethicists, social scientists, and legal scholars (particularly in the field of therapeutic jurisprudence).

Mental health professionals should note that the deference placed upon their judgment creates a heavy ethical as well as legal responsibility. They have an opportunity to achieve reform while advancing the interests of their patients, if they are prepared to stand by their professional views in the face of the corporate interests of their employer (or of those who are funding care services). It is interesting to note that right to treatment litigation arose in a context of large asylum care where psychiatrists, unless they were the owners or managers, were relatively powerless to influence the care given to patients.

Managed Care. The managed care context again severely limits professional autonomy, though now in hospital, clinic or community care settings, and raises vexing new ethical problems (Pettifor, 1996; Sederer and Bennett, 1996). As McCullough noted, "more than two decades of efforts to criticize — indeed, to discredit — paternalism threaten to become undone by the emergence of managed care" (1994, p. 940).

Will professionals be prepared to compromise care by trading off higher constraints for more income[44], or will they support reasonable litigation seeking further care and services, as well as critically evaluate the implications of corporate care for the health of their patients? Interestingly, as the health systems of most western countries undergo radical restructuring involving privatization and the formation of mega health management companies, both practitioners and

---

[43] For a discussion of systems ethics in health care see Strijbos (1994). See McCubbin and Cohen (1997a) for a discussion, grounded in ethics, of the implications of systems thinking for the study of mental health policy and politics.

[44] Such willingness seems to be evident in a recent survey of American psychiatrists by Schlesinger, Dorwart and Epstein (1996).

patients are finding that they share common ground over loss of their respective autonomy.

Affirmative Action. If society is truly concerned about formulating legal rights affecting mental patients, a concern based upon respect for their essential dignity and autonomy as persons, then it would be morally consistent to go much further in listening to them and their own advocates. This requires affirmative action, given the many disadvantages users bear with respect to making their views respected and influential in clinical, service, and court environments, and, more fundamentally, at the political level (Farber, 1995; McCubbin and Cohen, 1996). However, it is essential to be clear on the theory as to how an action may be "affirmative" as opposed to paternalistic, empowering as opposed to perpetuating dependence.

Especially at the present historical juncture, where pursuit in the courts of constitutional rights has gone a long way but has made little further progress over the past 15 years, at both the policy and service levels it is probably less pertinent to discuss which rights to paternalistically "give" patients than to reconstruct the policy and political playing fields to allow users themselves the means to assert their rights (McCubbin and Cohen, 1996). This means real participation on the boards governing health and social services and hospital authorities, and secure funding for user advocacy organizations (groups dominated by non-users concerned with "patient" interests are already well-funded). Without the tools for civic participation, the value for users of marginal rights gains can become quickly eroded as the health system restructures itself at an increasing rate.

The Autonomy Basis of the Right to Treatment. The concept of "right to treatment" served a useful purpose in a social context in which society was not ready to delve directly into the contradiction between the need for help of vulnerable persons and the autonomy interests of those same persons. An enlightened conception of parens patriae helps to untangle this thorny thicket: the State's obligation to "help" patients means to aid them in their own empowerment: to facilitate the development of their own powers.

Understood this way, the "right to treatment" becomes consequential to the already well-established human right of autonomy. Yet in recent years discussion of the "right to treatment" has been dominated by a profoundly paternalistic discourse concerned with mandatory treatment. Whether or not one shares the arguments for forced treatment, it is profoundly dehumanizing to those who receive such treatments to characterize interventions designed by others to be for the good of the patient, as, rather, responding to that patient's "rights". Such banalization of the meaning of rights can only augment growing societal cynicism which is tainting efforts to incorporate ethical principles into public policies.

References

Albee, G. W. (1968). Models, myths, and manpower. Mental Hygiene, 52, 168-180.

Appelbaum, P.S. (1987). Resurrecting the right to treatment. Hospital and Community Psychiatry, 38, 703-704, 721.

Aviram, U. (1990). Care or convenience? On the medical-bureaucratic model of commitment of the mentally ill. International Journal of Law and Psychiatry, 13, 163-177. P. 175.

Aviram, U. (1993). Screening services in civil commitment of the mentally ill: An attempt to balance individual liberties with needs for treatment. Bulletin of the American Academy of Psychiatry and Law, 21, 195-211.

Aviram, U., Minsky, S., Smoyak, S.A., and Gubman-Riesser, G.D. (1995). Discharge-ready patients who remain hospitalized: A re-emerging problem for mental health services. Psychiatric Quarterly, 66, 63-85.

Aviram, U., and Smoyak, S.A. (1994) Discharged pending placement: How courts created a new intermediate legal status for confining mentally ill persons. International Journal of Law and Psychiatry, 17, 139-151.

Aviram, U., and Weyer, R.A. (1996). Changing trends in mental health legislation: Anatomy of reforming a civil commitment law. Journal of Health Politics, Policy and Law, 21, 771-805.

Bazelon, D.L. (1969). The right to treatment: The court's role. Hospital and Community Psychiatry, 20, 129-135.

Beauchamp, T.L. and Childress, J.F. (1989). Principles of biomedical ethics (3rd ed.). New York: Oxford University Press.

Beck, J.C., and Parry, J.W. (1992). Incompetence, treatment refusal, and hospitalization. Bulletin of the American Academy of Psychiatry and Law, 20, 261-267.

Bell, N.K. (1996). Responsibilities and rights in the promotion of health: Differing positions of the individual and the State. Social Science and Medicine, 43, 775-782.

Bertalanffy, L. von (1981). A systems view of man. Boulder, Colorado: Westview.

Bertolote, J.M., and Sartorius, N. (1996). WHO initiative of support to people disabled by mental illness: Some issues and concepts related to rehabilitation. European Psychiatry, 11(Supp. 2), 56s-59s.

Birnbaum, M. (1960). The right to treatment. American Bar Association Journal, 46, 499-505.

Birnbaum, M. (1969). A rationale for the right. In D.S. Burris (Ed.), The Right to Treatment, 77-106. New York: Springer.

Birnbaum, M. (1974). The right to treatment: Some comments on its development. In F.J. Ayd (Ed.), Medical, moral and legal issues in mental health care, pp. 97-141. Baltimore: Williams and Wilkins.

Brown, P. (1985). The mental patients' rights movement, and mental health institutional change. In P. Brown (Ed.), Mental Health Care and Social Policy, 187-212. Boston: Routledge and Kegan Paul.

Burris, D.S. (Ed.). (1969). The right to treatment: A symposium. New York: Springer.

Chambers, D.L. (1972). Alternatives to civil commitment of the mentally ill: Practical guides and constitutional imperatives. Michigan Law Review, 70, 1107-1200.

Charest, A. (1994, 4 mai). Psychiatrie: Nouveau Code civil, nouvelles règles. L'Actualité médicale, 2.

Cohen, D., and McCubbin, M. (1990). The political economy of tardive dyskinesia: Asymmetries in power and responsibility. Journal of Mind and Behavior, 11, 465-488.

Cohen, D., Thomas, G., Dallaire, B., McCubbin, M., and Morin, P. (1997). Savoir, pouvoir, et dangerosité civile: Une étude quantitative et qualitative des décisions de révision de cure fermée de la Commission des affaires sociales du Québec, 1975-1993. Unpublished.

Council of the American Psychiatric Association (1967). Position statement on the question of adequacy of treatment. American Journal of Psychiatry, 123, 1458-1460.

Derdeyn, A.P. (1977). Child abuse and neglect: The rights of parents and the needs of their children. American Journal of Orthopsychiatry, 47, 377-387.

Dix, G.E. (1976). The Alabama "right to treatment" case: An opportunity not to be missed. Community Mental Health Journal, 12, 161-167.

Etzioni, A. (1988). The moral dimension: Toward a new economics. New York: Free Press.

Everett, B. (1994). Something is happening: The contemporary consumer and psychiatric survivor movement in historical context. Journal of Mind and Behavior, 15, 55-70.

Farber, S. (1995). Undermining community. Telos, 102, 177-192.

Feher, M. (1996). Empowerment hazards: Affirmative action, recovery psychology, and identity politics. Representations, 55, 84-91.

Foucault, M. (1962). Maladie mentale et psychologie. Paris: Presses universitaires de France.

Frank, R.H. (1985). Choosing the right pond: Human behavior and the quest for status. New York: Oxford University Press.

Froese, A.P. (1991). Minors' right to psychiatric treatment. Canadian Journal of Psychiatry, 36, 452-455.

Fromm, E. (1947). Man for himself: An inquiry into the psychology of ethics. Greenwich, Conn.: Fawcett.

Glendon, M.A. (1991). Rights talk: The impoverishment of political discourse. Toronto: Maxwell Macmillan.

Gunn, A.E. (1991). Risk-benefit ratio: The soft underbelly of patient autonomy. Issues in Law and Medicine, 7, 139-153.

Halpern, C.R. (1969). A practicing lawyer views the right to treatment. In D.S. Burris (Ed.), The right to treatment, 107-142. New York: Springer.

Hayes, R., Boerma, B., and Ovadia, F.T. (1993-94). Mental health and the law: An introduction. International Journal of Mental Health Services, 22(4), 3-21.

Hegel, G.W.F. (1952). The philosophy of right. The philosophy of history. Toronto: Encyclopaedia Britannica.

Hermann, D.H.J. (1990). Autonomy, self determination, the right of involuntarily committed persons to refuse treatment, and the use of substituted judgment in medication decisions involving incompetent persons. International Journal of Law and Psychiatry, 13, 361-385.

Isohanni, M., Nieminen, P., Moring, J., Pylkkänen, K., and Spalding, M. (1991). The dilemma of civil rights versus the right to treatment: Questionable involuntary admissions to a mental hospital. Acta Psychiatrica Scandinavica, 83, 256-261.

Jennings, A. (1994). On being invisible in the mental health system. Journal of Mental Health Administration, 21, 374-387.

Kant, I. (1958). Critique of pure reason (trans. and intro. by N.K. Smith; abridged). New York: Wiley.

Kapp, M.B. (1994). Treatment and refusal rights in mental health: Therapeutic justice and clinical accommodation. American Journal of Orthopsychiatry, 64, 223-234.

Kaufman, E. (1979). The right to treatment suit as an agent for change. American Journal of Psychiatry, 136, 1428-1432.

Kittrie, N.N. (1971). The right to be different: Deviance and enforced therapy. Baltimore: Johns Hopkins University Press.

Kleinman, I. (1986). Force-feeding: The physician's dilemma. Canadian Journal of Psychiatry, 31, 313-316.

Kunjukrishnan, R., and Varan, L.R. (1989). Interface between mental subnormality and law — A review. Psychiatric Journal of the University of Ottawa, 14, 439-452.

Lamb, H.R., and Mills, M.J. (1986). Needed changes in law and procedure for the chronically mentally ill. Hospital and Community Psychiatry, 37, 475-480.

Larner, G. (1995). The real as illusion: Deconstructing power in family therapy. Journal of Family Therapy, 17, 191-217.

Leaf, P. (1977). Wyatt v. Stickney: Assessing the impact in Alabama. Hospital and Community Psychiatry, 28, 351-356.

Locke, J. (1963). Two treatises of government (rev. ed.; intro. and notes by P. Laslett). Scarborough, Ontario: Mentor.

Loewy, E.H. (1993). Freedom and community: The ethics of interdependence. Albany, NY: State University of New York Press.

Maslow, A.H. (1970). Motivation and Personality (2nd ed.). New York: Harper and Row.

McCubbin, M. (1994). Deinstitutionalization: The illusion of disillusion. Journal of Mind and Behavior, 15, 35-54.

McCubbin, M. (1997, June). Substituted judgment: Problems in estimating the "pure best interests" of psychiatric patients considered incompetent. Presentation to the XXIInd International Congress on Law and Mental Health, International Academy of Law and Mental Health, Montréal.

McCubbin, M., and Cohen, D. (1996). Extremely unbalanced: Interest divergence and power disparities between clients and psychiatry. International Journal of Law and Psychiatry, 19, 1-25.

McCubbin, M., and Cohen, D. (1997a). Mental health policy from a systemic perspective: Orientations for strategic reform. Submitted.

McCubbin, M., and Cohen, D. (1997b). Les droits des usagers de services en santé mentale: Le noeud étroit du pouvoir, de la loi et de l'éthique. Submitted.

McCubbin, M., and Weisstub, D.N. (1998). Toward a pure best interests model of proxy decision making for incompetent psychiatric patients. International Journal of Law and Psychiatry, 21, 1-30.

McCullough, L.B. (1994). Should we create a health care system in the United States? Journal of Medicine and Philosophy, 19, 483-490.

McGough, L.S., and Carmichael, W.C. II (1977). The right to treatment and the right to refuse treatment. American Journal of Orthopsychiatry, 47, 307-320.

Mechanic, D. (1994). Establishing mental health priorities. Milbank Quarterly, 72, 501-514.

Miller, R.D. (1991). Law, psychiatry and rights. Medicine and Law, 10, 327-333.

Moorhouse, A., and Weisstub, D.N. (1996). Advance directives for research: Ethical problems and responses. International Journal of Law and Psychiatry, 19, 107-141.

Pellegrino, E.D. (1978). The social ethics of primary care: The relationship between a human need and an obligation of society. Mount Sinai Journal of Medicine, 45, 593-601.

Perlin, M.L. (1990). Are courts competent to decide competency questions? Stripping the facade from United States v. Charters. Kansas Law Review, 38, 957-1001.

Perlin, M.L. (1994). Law and the delivery of mental health services in the community. American Journal of Orthopsychiatry, 64, 194-208.

Perlin, M.L., Gould, K.K., and Dorfman, D.A. (1995). Therapeutic jurisprudence and the civil rights of institutionalized mentally disabled persons: Hopeless oxymoron or path to redemption? Psychology, Public Policy, and Law, 1, 80-119.

Petrila, J. (1992). Redefining mental health law: Thoughts on a new agenda. Law and Human Behavior, 16, 89-106.

Pettifor, J.L. (1996). Ethics: Virtue and politics in the science and practice of psychology. Canadian Psychology/Psychologie canadienne, 37, 1-12.

Rachlin, S. (1988). Litigating a right to treatment: Woe is me. Psychiatric Quarterly, 59, 182-192.

Rappeport, J.R. (1987). Belegaled: Mental health and the law in the United States. Canadian Journal of Psychiatry, 32, 719-727.

Rawls, J. (1971). A theory of justice. Cambridge, MA: Belknap.

Rosenau, P.V. (1994). Health politics meets post-modernism: Its meaning and implications for community health organizing. Journal of Health Politics, Policy and Law, 19, 303-333.

Rosenthal, E., and Rubenstein, L.S. (1993). International human rights advocacy under the "Principles for the Protection of Persons with Mental Illness". International Journal of Law and Psychiatry, 16, 257-300.

Rothman, D.J. (1980). Conscience and convenience: The asylum and its alternatives in progressive America. Glenview, Illinois: Scott, Foresman.

Rothman, D.J. (1991). Strangers at the bedside: A history of how law and bioethics transformed medical decision making. New York: BasicBooks.

Rousseau, J.-J. (1992). Du contrat social. Paris: Flammarion.

Schlesinger, M., Dorwart, R.A., and Epstein, S. (1996). Managed care constraints on psychiatrists' hospital practices: Bargaining power and professional autonomy. American Journal of Psychiatry, 153, 256-260.

Schopp, R.F. (1993). Therapeutic jurisprudence and conflicts among values in mental health law. Behavioral Sciences and the Law, 11, 31-45.

Schultz, G.S. (1996). Taxonomy of rights: A proposed classification system of rights for individuals with mental retardation or developmental disabilities. Journal of Developmental and Physical Disabilities, 8, 275-285.

Schwitzgebel, R. K. (1973). Right to treatment for the mentally disabled: The need for realistic standards and objective criteria. Harvard Civil Rights Law Review, 8, 513-535.

Sederer, L.I., and Bennett, M.J. (1996). Managed mental health care in the United States: A status report. Administration and Policy in Mental Health, 23, 289-306.

Seicshnaydre, S.E. (1992). Community mental health treatment for the mentally ill — When does less restrictive treatment become a right? Tulane Law Review, 66, 1971-2001.

Stambolovic´, V. (1996). Human rights and health within the dominant paradigm. Social Science and Medicine, 42, 301-303.

Stefan, S. (1992). Leaving civil rights to the "experts": From deference to abdication under the professional judgment standard. Yale Law Journal, 102, 639-717.

Stone, A.A. (1975). The right to treatment: Comments on the law and its impact. American Journal of Psychiatry, 132, 1125-1134.

Strijbos, S. (1994). The individual and the collective in health care: A problem of systems ethics. Systems Research, 11, 67-75.

Szasz, T. S. (1969). The right to health. In D.S. Burris (Ed.), The Right to Treatment, 59-76. New York: Springer.

Treffert, D.A. (1970). Alternatives to legislating the right to treatment. Hospital and Community Psychiatry, 21, 302-303.

Twerski, A.D. (1971). Treating the untreatable: A critique of the proposed right to treatment law. Hospital and Community Psychiatry, 22, 261-264.

Ubel, P.A., and Loewenstein, G. (1996). Distributing scarce livers: the moral reasoning of the general public. Social Science and Medicine, 42, 1049-1055.

Wald, P.M., and Friedman, P.R. (1978). The politics of mental health advocacy in the United States. International Journal of Law and Psychiatry, 1, 137-152.

Weitzel, W.D. (1977). Changing law and clinical dilemmas. American Journal of Psychiatry, 134, 293-295.

Wilk, R.J. (1994). Are the rights of people with mental illness still important? Social Work, 39, 167-175.

Winick, B.J. (Ed.). (1996). A critical examination of the MacArthur Treatment Competence Study: Methodological issues, legal implications, and future directions. Psychology, Public Policy, and Law, 2(1).

Wolf, S.M. (1994). Health care reform and the future of physician ethics. Hastings Center Report, 24(2), 28-41.