

2

Méthodologie

LES STRATÉGIES DE LA RECHERCHE SPÉCULATIVE

- 1 La constitution du corpus des énoncés de base
 - 1.1 *Le corpus unique, intertextuel ou contrasté*
 - 1.2 *La qualité et la validité des corpus*
- 2 L'analyse conceptuelle
 - 2.1 *Les objectifs de l'analyse conceptuelle*
 - 2.2 *Les perspectives et les niveaux d'analyse*
 - 2.3 *Les méthodes de l'analyse conceptuelle*
- 3 L'analyse critique
 - 3.1 *Les objectifs de l'analyse critique*
 - 3.2 *Les tactiques de l'analyse critique*
 - 3.3 *Les pièges de l'analyse critique*
- 4 L'analyse inférentielle
 - 4.1 *Les objectifs de l'analyse inférentielle*
 - 4.2 *Les méthodes de l'analyse inférentielle*
 - 4.3 *Les pièges de l'analyse inférentielle*
- 5 Les questions stratégiques
 - 5.1 *Les fausses évidences de l'analyse du manque*
 - 5.2 *De quelques stratégies de conviction*
 - 5.2.1 L'argument étymologique
 - 5.2.2 L'argument historique
 - 5.2.3 L'argument du bon sens et du sens commun
 - 5.2.4 L'argument anecdotique
 - 5.2.5 L'argument par le jeu de mots
 - 5.2.6 L'argument de la logique et de la raison
 - 5.2.7 L'argument de l'éthique et de la métaphysique

La recherche spéculative a été définie comme un travail de l'esprit produisant des énoncés théoriques à partir et à propos d'autres énoncés théoriques. Comme la recherche empirique, elle vise la théorisation, mais elle s'en différencie dans la mesure où elle ne travaille pas sur des données empiriques, sur des traces ou des inscriptions qu'elle aurait produites pour appuyer son argumentation. Cependant, la théorisation ne peut se faire à partir de rien : même la recherche spéculative a besoin d'un matériel, soit au moins de quelques autres énoncés théoriques produits antérieurement ou ailleurs. Dès lors, avant d'examiner trois des formes de l'analyse spéculative, il faut envisager la question du matériel à partir duquel la théorisation s'opère; c'est le problème de la constitution préalable du corpus des énoncés de base.

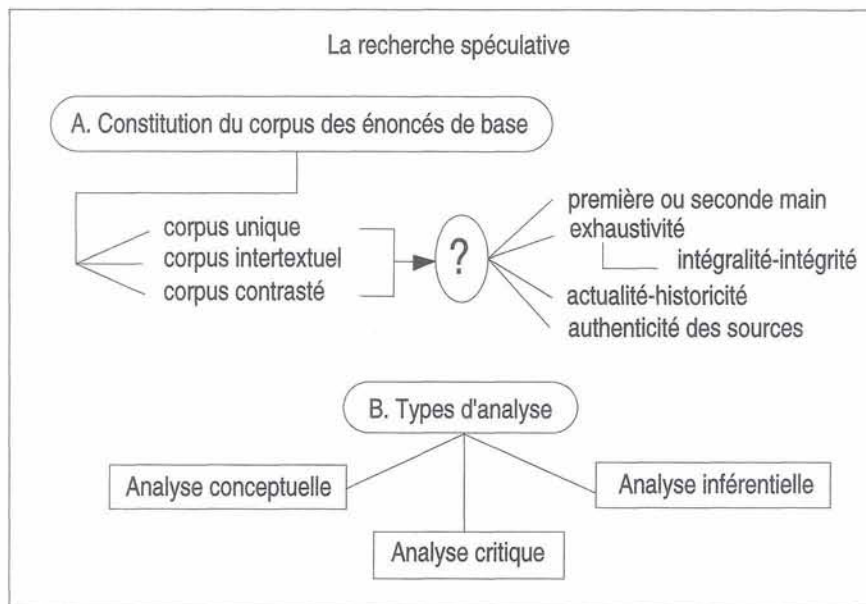


Figure 6.1 : La recherche spéculative. Plan général.

Notons, en passant, que toutes les recherches empiriques, qu'elles soient nomothétiques ou appliquées, comportent quasi nécessairement une partie de recherche spéculative et que les réflexions qui suivent s'appliquent à ce l'on envisagera plus loin (chapitre 17) sous l'appellation de *cadre conceptuel* et de *cadre théorique*. À moins de n'être qu'une simple compilation, les cadres conceptuels et théoriques ne peuvent être produits que par une recherche spéculative, ce qui constitue une première difficulté pour le chercheur en formation.

1 La constitution du corpus des énoncés de base

La première tâche d'une recherche spéculative consiste à sélectionner les énoncés théoriques à partir desquels le travail de réflexion se bâtira. Habituellement, deux situations sont possibles : ou bien le chercheur débute dans le domaine et il doit s'appropriier l'ensemble des écrits à partir desquels il va travailler, ou bien il s'agit d'un chercheur expérimenté, auquel cas il extraira de ses connaissances les énoncés qui lui sembleront pertinents.

1.1 *Le corpus unique, intertextuel ou contrasté*

Si l'objet de la spéculation porte sur l'interprétation des énoncés d'un discours théorique à propos d'une notion (par exemple, la signification de la coopération dans la pédagogie de Freinet), un *corpus unique* suffira : il s'agira alors de réunir les écrits de cet auteur et d'en extraire les passages se rapportant au sujet de l'étude. Encore que cet auteur ait pu tenir des discours sur le même sujet mais à partir de points de vue différents, soit en s'adressant tantôt à des élèves, tantôt à des parents, à des politiciens, à des administrateurs, en situation de paix sociale ou en situation de conflit; dans ce cas, on se trouve probablement devant la nécessité de constituer un corpus *intertextuel*. Dans le cas où les chercheurs travaillent sur un concept, une notion, une idée, ils s'appuient sur des énoncés provenant de plusieurs discours et l'on parlera alors de corpus intertextuel ou de corpus contrasté.

Constituer un *corpus intertextuel*¹ revient non seulement à chercher des énoncés produits par plusieurs auteurs sur un thème ou par un seul auteur s'adressant à des lecteurs différents dans des situations variées, mais aussi à identifier les différences dans les conditions de production de ces énoncés. On parle de corpus intertextuel dans la mesure où le contexte de l'énonciation, tant du point de vue de l'émetteur que du point de vue des récepteurs réels ou envisagés, est considéré comme déterminant aussi bien la forme que le contenu des énoncés. Dans une perspective intertextuelle, le sens de l'énoncé se construit à l'intersection des conditions de l'énonciation. Dès lors, l'interprétation juste et nuancée d'un énoncé ne sera possible que par l'examen concomitant des différences entre les formes des énoncés et des différences entre les contextes. Par exemple, on ne peut faire une analyse des discours du ministre de l'éducation sur le financement

1 Nous reprenons cette notion et la méthodologie qui l'accompagne à D. Jacobi, *Textes et images de la vulgarisation scientifique*. Berne : Peter Lang, 1987. Voir pages 74 à 89 : «Intertexte et dialogisme».

des universités et sur les frais de scolarité que dans la mesure où l'on tient compte de ses auditeurs (étudiants, professeurs d'université, recteurs, électeurs payeurs de taxes), de la proximité d'une période électorale, de l'état variable d'endettement de chacune des universités, etc.

Un *corpus contrasté* est constitué d'énoncés provenant d'auteurs qui ont des options, des préconceptions, des points de vue différents à propos d'une notion ou d'un événement. Les énoncés sont alors recherchés parce que leur différence permet de mieux voir l'éclatement ou les variations des discours. Si le corpus intertextuel permet de dégager ce qui est commun, ce qui est partagé par les énoncés au delà des variations contextuelles, et s'il permet aussi de mettre en évidence la souplesse des formulations autour du noyau notionnel, le corpus contrasté permet d'éclairer la richesse des écarts, des tensions entre les énoncés : il permet de se faire une représentation de l'éventail des significations. Pour prendre une image empruntée à la statistique, on dira que le corpus intertextuel est sensible aux effets de tendance centrale, tandis que le corpus contrasté est plus sensible aux effets de dispersion. Le choix d'un corpus unique, intertextuel ou contrasté dépend donc non seulement de l'objet de la recherche spéculative, mais aussi de sa perspective; on voit là l'importance que prend l'explicitation préliminaire de ses choix par le chercheur.

1.2 La qualité et la validité des corpus

Quel que soit le type de corpus envisagé, la crédibilité de la recherche spéculative reposera aussi sur la qualité et la validité du corpus constitué. Quatre critères sont envisagés : l'accès aux sources, l'exhaustivité, l'actualité et l'authenticité.

Un premier critère de crédibilité du corpus réside dans le mode d'*accès aux sources* : s'agit-il d'un corpus de *première main* ou de *seconde main* ? Les énoncés extraits directement des écrits d'un auteur ont une tout autre valeur que les énoncés de cet auteur rapportés par un autre. Dans ce dernier cas, plusieurs questions doivent être résolues : s'agit-il d'une citation exacte, d'une citation tronquée et reconstruite, d'une reformulation synthétique ou d'une interprétation ? Si la vérification à la source d'une citation de seconde main ne peut être faite, le chercheur donnera plus de crédibilité à l'énoncé qu'il sélectionne pour son corpus s'il se contente de l'attribuer à la seconde main dont il a lu le texte, plutôt qu'au premier dont il n'est pas sûr que les écrits ont été fidèlement rapportés. Le problème de l'accès à la première source est d'autant plus important qu'il est probable

que le second auteur ne cite le premier que pour appuyer, pour renforcer ses propres énoncés. Dès lors, il est plus que probable qu'il mette l'accent seulement sur ce qui lui convient : la copie risque d'être un simulacre qui enfle l'appui à ses énoncés en omettant les nuances, ou une caricature qui grossit les manques, les défauts ou les oppositions pour mieux se qualifier.

La prudence préconise donc de bien distinguer les citations de première main et les citations de seconde dont on ne peut que difficilement connaître le statut. Cette prudence se justifie d'autant plus que le recours aux citations de seconde main n'aboutit souvent qu'à donner une illusion d'érudition en cette fin de XX^e siècle. En effet, si l'on pouvait encore au milieu du siècle prétendre faire le tour des écrits sur un sujet, cela n'est plus possible aujourd'hui, à une époque où les chercheurs, pour survivre, doivent souvent beaucoup trop écrire. Lorsque l'on travaille à partir de travaux de synthèse, il faut citer directement les auteurs de ces synthèses et leur laisser la responsabilité de l'interprétation des auteurs qu'ils ont condensés, sans chercher à donner l'illusion d'une grande érudition en allongeant la bibliographie de toutes les citations de seconde et d'énième mains. Comme en cuisine, où une sauce trop diluée est fade, une bibliographie gonflée perd de sa valeur et de sa crédibilité. Surtout si on fait dire aux auteurs des choses qu'ils n'ont jamais écrites !

La sélection des extraits pose le problème de l'*exhaustivité et de l'intégralité* (y compris l'intégrité) des textes cités par rapport à ceux qui auraient pu ou auraient dû l'être. Comme il est impossible et inutile que la réflexion porte sur la totalité et l'intégralité des textes qui touchent au problème posé, le chercheur doit sélectionner les extraits qui se rapportent au problème étudié : l'exhaustivité est donc relative et non pas absolue. Deux problèmes se posent alors.

1° Le chercheur devrait sélectionner tous les extraits pertinents, y compris ceux qui ne vont pas dans le sens de ses anticipations, y compris ceux qui comportent des difficultés, des contradictions, des nuances, des paradoxes. Il ne s'agit pas de nettoyer le corpus en éliminant ce qui pose problème ou ce qui entre en contradiction avec les idées du chercheur.

2° Dans sa sélection, le chercheur doit garder ses extraits complets. Citant une phrase complexe, le chercheur ne peut en gommer certaines propositions qui nuancent, qui relativisent ou remplacent certains mots par d'autres qui rendraient la lecture plus simple, plus conforme aux hypothèses anticipées. De plus, le sens de certaines phrases change lorsqu'elles sont

isolées du *cotexte*², c'est-à-dire des phrases qui les entourent. Il importe donc que le chercheur conserve le cotexte des extraits, si une variation du sens peut se produire lorsqu'on les isole. Cela est particulièrement important lorsque les extraits sont repris à des textes techniques qui, comme c'est souvent le cas en sciences humaines, utilisent des mots du langage courant en leur attribuant un sens technique bien particulier. L'erreur fréquente consiste alors à remplacer le sens technique par le sens commun, ce qui permet d'appuyer, par un discours scientifique malmené, une déclaration que ce discours n'appuie pas. Autant l'interprétation hors contexte d'un énoncé peut conduire à des méprises, autant l'utilisation hors cotexte de citations techniques risque de produire des impostures.

Un autre problème fréquent dans une discipline appliquée comme l'éducation réside dans l'*actualité* et l'*historicité* des sources, selon le cas.

Le chercheur qui compose un corpus se rapportant à une notion, doit recueillir les extraits représentatifs de l'état actuel du discours se rapportant à cette notion, et non pas se contenter de ce qui s'en disait il y a vingt et quelques années. Or, lorsqu'il s'agit de notions qui se sont développées d'abord dans les disciplines contributives, les chercheurs en éducation ont tendance à ne pas être au courant des travaux contemporains s'y rapportant dans la discipline contributive. Pour aborder la notion en question, ils privilégient une approche par textes de vulgarisation spécialisée, par traités ou par manuels d'enseignement, ouvrages qui, étant donné les délais nécessaires à leur composition et le public qu'ils visent, ne donnent qu'une image passée et souvent simplifiée du discours sur la notion. Si ce genre de lectures est un passage nécessaire avant d'aborder les travaux contemporains (articles de revues, parfois chapitres d'ouvrages à auteurs multiples et publications originales), le chercheur ne peut s'y arrêter et, sans négliger le développement historique de la notion, son matériel doit être actuel et refléter les écrits contemporains.

Lorsque les écrits se rapportent au passé (proche ou lointain) ou lorsque les écrits relevés sont datés, l'authenticité des sources est à surveiller. Cela correspond au problème déjà discuté de l'accès en première ou en seconde main. D'autre part, la datation des extraits est importante, car elle

2 On distinguera entre le cotexte, c'est-à-dire l'ensemble des éléments graphiques (texte, illustrations, titres, etc.) qui entourent le texte, et le contexte, c'est-à-dire les conditions (environnement historique, socioculturel, relationnel, etc.) dans lesquelles le texte a été produit. Cette distinction est reprise à D. Jacobi, «Le discours de vulgarisation scientifique, problèmes sémiotiques et textuels», dans D. Jacobi et B. Schiele, *Vulgariser la science*, Paris-Seysse : Champ Vallon-PUF, 1988. Voir la note 9 pages 100 et 115.

donne une idée du contexte de l'écrit et de sa filiation. Un texte écrit au moment d'un événement, par un témoin y ayant joué un rôle, apporte nécessairement une autre information que le texte écrit par le même témoin, mais après réflexion, une fois les choses terminées; l'anecdote est probablement plus exacte, plus riche dans le premier cas, alors que le sens de l'événement est probablement plus ancré et contextualisé dans le second. Lorsqu'il s'agit de textes se rapportant à une notion, leur datation peut indiquer le cheminement des pensées, la dépendance des auteurs successifs et les problèmes pratiques, théoriques ou politiques auxquels les écrits peuvent être une réaction. Par exemple, quasi tous les textes se rapportant aux plans expérimentaux après 1963 reprennent les graphiques de l'article original de Campbell et Stanley, y compris les symboles utilisés. L'examen des références et de leurs dates donne des indications sur l'école de pensée des auteurs et sur leur honnêteté intellectuelle.

La constitution d'un corpus d'énoncés de base pose donc des problèmes semblables à ceux envisagés par la critique du témoignage dans la recherche historique (ou policière). Le type de corpus ayant été choisi en fonction du problème de recherche, le corpus ayant été constitué et sa validité établie, trois modalités d'analyse sont à envisager, chacune ayant des possibilités différentes : soit les analyses conceptuelle, critique et inférentielle.

2 L'analyse conceptuelle

2.1 Les objectifs de l'analyse conceptuelle

L'analyse conceptuelle a pour objectif de dégager le sens et les possibilités d'application d'un concept ou d'une notion, en identifiant les constituants du champ sémantique de ce concept ou de cette notion et ses interactions avec d'autres champs. L'analyse conceptuelle tentera de dégager, par diverses comparaisons, quelle est l'intention ou la compréhension du concept et quelle est son extension ou son étendue. Par intention, on entend le noeud des significations qu'il peut porter, c'est-à-dire ce qu'il dit seulement et nécessairement, ou encore le sens strict, soit le sens réduit aux constituants communs à l'ensemble de ses utilisations. Par étendue ou extension, on entend ce qu'un concept peut dire de plus lorsqu'il est utilisé dans certaines situations; il s'agit alors d'examiner quelles sont les différentes situations (ou les sujets) auxquelles ce concept s'applique, à quelles conditions il peut s'y appliquer et quelles nuances (ou modalités de sens) cette application introduit par rapport à l'intention. L'analyse conceptuelle

est une base de la théorisation, puisque, ainsi comprise, elle permet de mieux cerner ce qu'un concept veut dire, à quels sujets il peut s'appliquer et à quelles significations ces applications conduisent.

Le préalable à l'analyse conceptuelle consiste à délimiter le concept théorique problème que l'on souhaite examiner. Certains concepts techniques ne posent guère de problème parce que leur définition est précise et que leur sens est restreint à une opérationnalisation formelle comme une formule mathématique. Autrement dit, l'analyse conceptuelle manque d'intérêt pour les concepts dont l'extension est réduite à l'intention : on ne passera pas son temps à faire l'analyse conceptuelle de la moyenne arithmétique. Par ailleurs, les préconceptions qu'en ont les élèves pourraient faire l'objet d'une analyse, mais alors il s'agit d'une étude d'épistémologie génétique et pas d'une analyse conceptuelle. D'autres concepts sont très problématiques; par exemple, en psychanalyse, le concept de pulsion est utilisé de tant de manières (10 entrées dans le *Vocabulaire* de Laplanche et Pontalis³) qu'il devient problématique de cerner son sens. Des mots comme validité, éducation permanente, formation continue, autonomie, sont utilisés dans une telle variété de contextes qu'il devient intéressant d'engager une analyse conceptuelle à leur propos.

2.2 Les perspectives et les niveaux de l'analyse

Les comparaisons de texte devraient se faire selon plusieurs perspectives : dans une perspective historique, évolutive, au niveau des discours théoriques, mais aussi au niveau de l'opérationnalisation, c'est-à-dire des procédures et des instrumentations concrètes retenues par les chercheurs pour produire des inscriptions valides des concepts théoriques.

La première phase, souvent oubliée bien qu'instructive, consiste à relever l'histoire du concept à l'étude, appelé *pivot*. On fait porter l'analyse sur des extraits représentatifs de l'évolution historique du concept. Il s'agit de retracer ses premières utilisations, les premiers auteurs, de façon à examiner la possibilité de transformations importantes entre son usage originel et les usages contemporains. Se pose alors la question du moment et du contexte dans lequel les transformations ou les glissements sémantiques se sont produits, afin d'identifier les événements ou les facteurs socioculturels auxquels ce concept est sensible.

3 Laplanche J. et J.B. Pontalis, *Vocabulaire de la psychanalyse*. Paris : PUF, 1971 (3e édition).

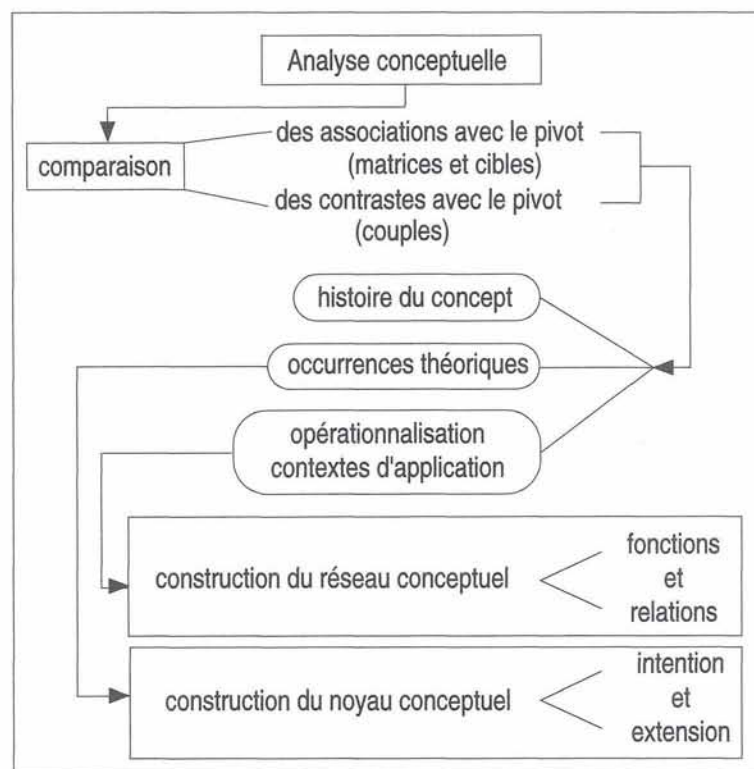


Figure 6.2 : L'analyse conceptuelle.

Ensuite, on effectue les comparaisons en examinant les relations entre les diverses occurrences théoriques⁴ du concept faites par différents auteurs ou par le même auteur dans différents contextes et dans différents contextes. Comme on l'illustrera plus loin dans un petit exemple à propos de la validité interne, les occurrences théoriques d'un concept sont parfois telles que ses significations peuvent être bien différentes et très larges.

Mais on rencontre encore des cas où les analyses faites sur les occurrences théoriques du concept semblent converger. Il est alors utile de voir quelles opérationnalisations les auteurs font de ce concept, c'est-à-dire d'analyser de manière comparative les associations du concept-cible avec les procédures décrites par les auteurs afin de le mesurer, de l'observer. C'est parfois en observant de quels indices les auteurs se servent pour retracer leurs concepts dans la réalité que l'on peut noter les différences

⁴ Occurrences théoriques : passages, extraits de discours théoriques dans lesquels le concept cible ou terme-pivot est utilisé.

qui élargissent le champ du concept-cible. L'examen des systèmes d'inscription utilisés pour cerner le concept permettra aussi d'établir la carte ou le réseau conceptuel. Pour ce faire on identifiera d'abord ses fonctions, c'est-à-dire à quoi peuvent servir les éléments qui opérationnalisent le concept dans la définition d'autres concepts. Ensuite on repère ses relations : quels éléments la définition du concept-pivot emprunte à d'autres concepts et quels éléments sont communs avec d'autres concepts. La comparaison des opérationnalisations du concept par différents auteurs peut s'élargir avec profit en examinant les différences d'opérationnalisation entre différentes disciplines ou sous-disciplines. Les problèmes de la « validité interne » que l'on va observer dans l'exemple ci-dessous provient peut-être des appartenances disciplinaires différentes des auteurs (psychologie, éducation, sociologie... dans le désordre!).

La mise en évidence de relations entre les composantes des différents extraits dans une perspective historique, au niveau des occurrences théoriques et des opérationnalisations, permettra alors d'organiser le champ conceptuel du terme pivot. Ce sera souvent la classification des similitudes et des différences d'association (analyses des cotextes du pivot), suivie de la mise en relation de ce classement avec les différents contextes de sa production, qui fourniront les intuitions permettant au chercheur de voir clair et de construire une représentation identifiant l'intention et l'étendue du concept analysé.

2.3 Les méthodes de l'analyse conceptuelle

L'examen des variations entre les utilisations du concept peut s'établir par la comparaison des associations entre ce concept cible ou terme-pivot et les autres mots (concepts, attributs, propositions) utilisés pour le définir en tenant compte de la proximité de chacune des associations. Toutes les méthodes d'analyse de contenu manifeste (voir chapitres 18 et 19), qu'elles soient intégrales ou sélectives, sont utilisables en analyse conceptuelle. Le choix d'une ou de plusieurs de ces méthodes dépendra de l'ampleur du matériel à analyser, de la finesse recherchée et du niveau visé par l'analyse conceptuelle. L'analyste sera souvent amené à faire des compromis car, la plupart du temps, le matériel est abondant et une analyse fine ne peut se faire sur tout le matériel. La rigueur peut être maintenue dans ces compromis en choisissant, sur la base d'une règle (voir chapitre 19, la détermination des unités d'analyse), les extraits auxquels on appliquera une analyse fine du texte (voir chapitre 18.6, les analyses du discours).

Illustration

Dans l'exemple qui suit, l'auteure a effectué une étude intertextuelle⁵ sur la notion de validité. Elle a comparé des extraits de manuels de méthodologie. Les phrases et les paragraphes comportant le terme-pivot «validité interne» étaient extraits et découpés en propositions, de telle sorte que l'on retienne au maximum les cinq propositions introduisant et les cinq propositions explicitant le pivot. Cette limite de cinq est arbitraire, mais une limite est nécessaire afin de pouvoir travailler sur des extraits manipulables. Une fois les extraits sélectionnés, l'auteure a construit des matrices dans lesquelles la succession des colonnes indique la proximité de l'association avec le pivot. Les matrices peuvent avoir des colonnes négatives dans lesquelles s'inscrivent les associations qui introduisent le concept-cible ou qui lui sont antérieures, et des colonnes positives où l'on place les associations qui explicitent ou qui suivent le concept-cible.

Pour l'illustration ci-dessous, on a retenu deux extraits (1a et 1b) produits par un premier auteur qui définit la validité interne. Le premier extrait est découpé en deux propositions qui se suivent : «La validité interne est la capacité..., elle correspond à la question...». Le terme pivot «validité interne» est situé au niveau 0, la proposition «capacité de l'expérience...» est située au niveau 1, et la proposition «correspond à la question...» est située au niveau 2. On a procédé de la même manière pour l'extrait 1b et pour les extraits repris aux auteurs numéro 2 et numéro 5. Cela donne le tableau de la page suivante.

Une exploitation superficielle de ce découpage décalé des propositions associées au pivot «validité interne», pour quatre extraits provenant de trois des auteurs examinés, met en évidence que l'auteur n° 5 ne parle pas de la même chose que les auteurs n° 1 et n° 2 : il parle plus de la fidélité des mesures que de la consistance programmatique. Si l'on examine les extraits des auteurs n° 1 et n° 2, on observe qu'ils parlent tous deux de consistance programmatique, mais que les extraits 1a et 1b insistent en colonne +1 sur la qualité de l'expérience, alors qu'au même niveau, le n° 2 parle d'exigence minimale pour interpréter les résultats. Si la nuance est subtile, elle est quand même là. L'auteur n° 1 se place plus que le n° 2 dans une perspective de vérification d'hypothèse théorique, alors que le n° 2 semble plutôt se situer au niveau de l'évaluation des effets d'un traitement soumis à une épreuve de contrôle par expérimentation.

5 Mainville S., *Dialogique de la validité : énoncés de manuels et de praticiens*. Mémoire de M.A. en sciences de l'éducation, Département d'études en éducation et d'administration de l'éducation, Université de Montréal, 1989.

Cette rapide analyse montre que le terme-pivot a deux sens différents chez les auteurs, ce qui n'est pas sans brouiller les cartes. En nous restreignant aux exemples ci-dessus (ils sont confirmés par d'autres extraits), on peut avancer l'hypothèse que le terme «validité interne» a deux sens (deux intentions), c.-à-d. celui de la fidélité des mesures (extrait 5) et celui de la consistance programmatique (extraits 1a, 1b et 2), et que cette seconde intention a au moins deux applications (deux extensions), celle de la mise à l'épreuve d'une hypothèse théorique par une expérimentation et celle de l'évaluation d'un traitement par une expérimentation. L'observation d'un double sens n'est pas rare dans la langue, mais elle pose problème lorsque les deux sens existent dans un même domaine : cela conduit à des contresens qui nous font penser que le sens de l'extrait 5 (bien que partagé par quelques autres) est une erreur.

De telles analyses peuvent se faire à la main comme elles peuvent s'effectuer par ordinateur, en compilant les fréquences avec lesquelles tel concept se retrouve associé à la cible (ou terme-pivot) dans une position d'ordre +1, +2, +3, +n, et -1, -2, -3, -n. Cependant la comparaison des profils (ou portraits) complets de chaque chaîne d'associations est souvent plus instructive que l'examen de la fréquence de chacune des associations : on apprend souvent plus en examinant les différentes listes d'associations de -n à +n, qu'en sachant combien de fois tel mot est associé à la cible dans une position +1 ou -1. De plus, si l'on considère chacun des extraits comme produit par une source (un auteur) et que chaque source peut être identifiée comme marquée par un trait spécifique (date, spécialité de l'auteur, public cible de l'écrit, contexte, etc.), on peut, de manière plus sophistiquée, effectuer une analyse factorielle des correspondances⁶. Cette analyse permet de situer graphiquement, sur un plan à deux dimensions, chacune des sources par rapport à l'ensemble des extraits, chacun des extraits par rapport à l'ensemble des sources, et chacun des extraits et des sources par rapport aux éléments des extraits qui pondèrent les deux principaux axes de l'analyse. Cette analyse permet alors de mettre en évidence les similitudes entre sources, les similitudes entre extraits et, de là, d'établir l'intention du concept et les situations qui délimitent chacune de ses extensions.

6 Lebart L. et A. Salem, *Analyse statistique des données textuelles*. Paris : Dunod, 1988.

3 L'analyse critique

3.1 Les objectifs de l'analyse critique

L'analyse critique a pour fin d'évaluer un ensemble d'énoncés théoriques afin de mettre en évidence ses lacunes, ses contradictions, ses paradoxes, ses conditions, ses présupposés, ses implications et ses conséquences, la plupart du temps non dites par les premiers auteurs. En effet, cette analyse critique a souvent comme objectif ou bien de condamner une théorie pour la remplacer par une autre, ou bien de proposer des améliorations, des reformulations, des compléments qui lui permettraient de paraître plus résistante sur le plan logique et plus crédible d'un point de vue psychologique et sociologique. Cet objectif indique la méthodologie de base : la comparaison avec une théorie formelle énonçant ce que la théorie «émissaire⁷» devrait être, ou avec une théorie considérée comme plus forte.

3.2 Les tactiques de l'analyse critique

Les phases de l'analyse critique peuvent se schématiser comme dans la figure 6.3. Une fois la théorie émissaire choisie, il faut d'abord examiner à quel type théorique elle appartient, pour pouvoir choisir la théorie étalon (ou de référence), c'est-à-dire pour préciser les exigences que l'on impose à la théorie qui sera examinée. Les exigences essentielles et leur pertinence devraient être argumentées en fonction du type de théorie dont relève la théorie qui sera analysée.

La seconde étape consisterait à repérer ses concepts les plus naïfs, ceux qui sont les plus faibles. Leur faiblesse sera examinée en fonction de quatre critères :

1° Parce qu'ils sont moins bien enchaînés aux autres concepts, qu'on ne voit pas quel lien les relie, qu'on soupçonne un saut inductif illicite;

2° Parce qu'ils ne répondent pas aux exigences essentielles imposées par le modèle étalon correspondant au type théorique auquel appartient la théorie dont ils font partie;

3° Parce qu'on soupçonne qu'ils ont des implications, des conséquences ou des présupposés non dévoilés qui pourraient introduire des contradictions ou des ruptures dans la théorie;

7 Émissaire : destiné à Azazel, le démon du désert. Comme le bouc émissaire qui porte tous les péchés du peuple et qui sera sacrifié à sa place.

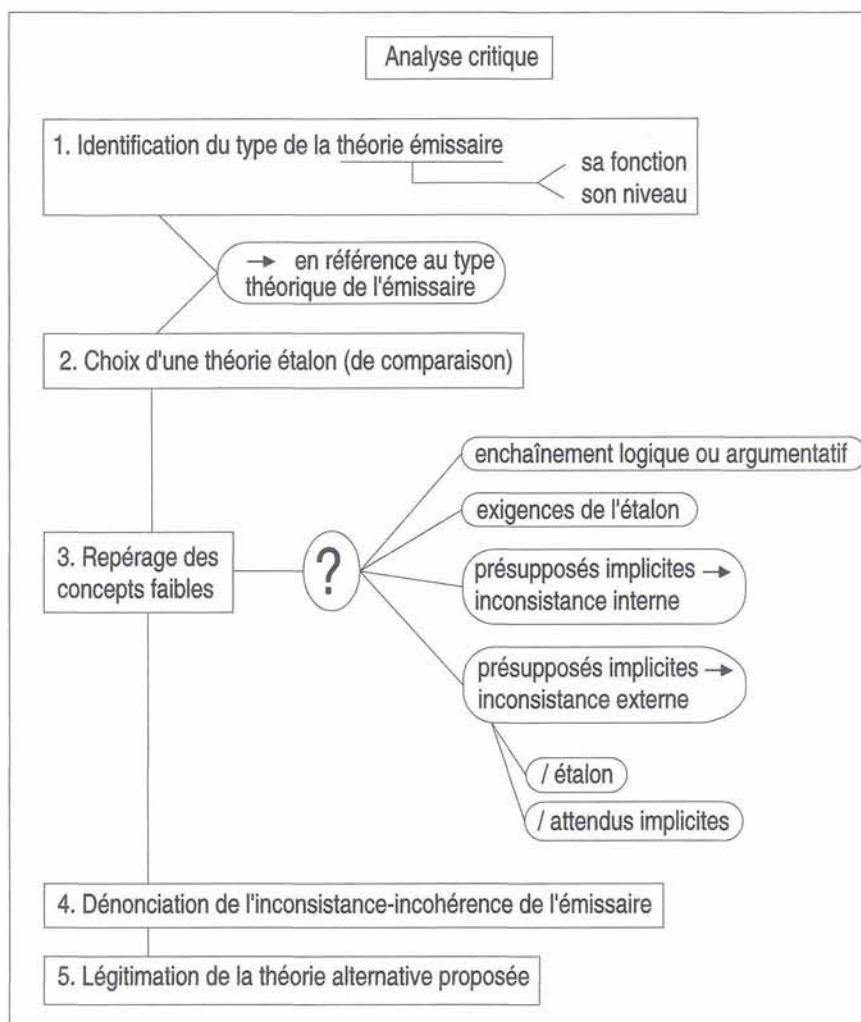


Figure 6.3 : L'analyse critique.

4° Parce qu'on soupçonne qu'ils ont des implications, des conséquences ou des présupposés non dévoilés qui, si on les dévoilait, pourraient être en opposition ou être inacceptables par rapport à ce qui était implicitement attendu ou perçu par les lecteurs de cette théorie.

La troisième phase consiste alors à proposer de nouveaux concepts en remplacement des concepts affaiblis et à montrer comment ces nouveaux concepts complètent avec consistance la théorie et comment ils respectent eux-mêmes le modèle étalon qui a permis la critique des concepts initiaux.

3.3 *Les pièges de l'analyse critique*

La pratique de l'analyse critique rencontre souvent quelques difficultés. D'abord, on observe que certains chercheurs inversent les phases : ils ne dévoilent pas d'abord le modèle étalon et ils commencent directement par la critique des concepts faibles. Or, dans certains cas leur modèle étalon est lui-même faible, car ils le construisent progressivement, de manière *ad hoc*, sans vérifier si ce modèle étalon est lui-même consistant et sans s'être assuré qu'il correspondra au type théorique une fois que l'analyse sera complétée. Le renversement des phases introduit donc une double faiblesse à éviter : la suspicion d'inconsistance et la suspicion d'impertinence.

Une seconde difficulté est souvent liée au fait que, par suite d'un soi-disant manque de place pour tout écrire, les extraits retenus pour présenter la théorie émissaire et le modèle étalon sont trop réduits, hors cotexte et caricaturaux. Il est en effet facile de supprimer d'un extrait les éléments du cotexte qui introduisent des nuances et des subtilités, pour confronter ce qui reste avec un modèle concis, idéal, face auquel l'émissaire sera condamné à l'avance.

Enfin, l'évidence de l'enrichissement apporté par les nouveaux concepts paraît parfois telle aux yeux du chercheur qu'il en oublie que les concepts qu'il apporte devront subir les mêmes critiques que celles qu'il a faites aux concepts naïfs de la théorie émissaire. C'est à lui de faire la démonstration de leur efficacité sur le plan de la consistance logique de la théorie remaniée autant que des opérationnalisations, et cela sans attendre que d'autres se donnent le plaisir d'en montrer les failles.

4 *L'analyse inférentielle*

4.1 *Les objectifs de l'analyse inférentielle*

L'analyse inférentielle a pour objectif le développement ou l'extension de théories. L'extension d'une théorie existante et de ses applications dans son champ originel se réalise par l'ajout d'éléments théoriques inférés. Le développement d'une nouvelle théorie dans un domaine donné s'obtient par le transfert d'une théorie d'un autre domaine à la suite de la perception d'une analogie entre les domaines. Dans les deux cas, il s'agit d'inférer de nouveaux éléments théoriques à partir d'éléments existants et, même lorsqu'il s'agit d'une extension interne, c'est la plupart

du temps à partir d'apports externes que l'extension peut se faire. L'utilisation du terme «extension» correspond à celle donnée plus haut : dans la majorité des cas, il s'agit bien d'adjoindre quelques éléments théoriques qui permettent d'appliquer la théorie à un domaine redélimité (soit d'en augmenter l'étendue, soit de la restreindre par des spécifications de conditions).

4.2 Les méthodes de l'analyse inférentielle

Deux cas sont à distinguer : l'extension interne d'une théorie et son extension à un autre domaine (ou son transfert).

Lorsqu'il s'agit de l'extension interne d'une théorie, l'analyse inférentielle procède en effectuant un examen de la signification, de l'enchaînement, des implications et des conséquences des concepts et de leurs opérationnalisations ainsi que de la pertinence de ces dernières. Cet examen se rapproche de l'analyse critique, mais avec une perspective de développement plutôt que de condamnation et de remplacement. On procède selon l'une ou plusieurs des stratégies suivantes (voir figure 6.4).

1. L'examen des enchaînements aux autres concepts, afin d'inférer des chaînons manquants qui déboucheraient sur de nouvelles applications ou sur des restrictions, ou afin de dégager les implications des faiblesses observées (ce qui conduit le plus souvent à des restrictions).

2. L'examen de la correspondance aux exigences essentielles imposées par le modèle étalon correspondant au type théorique auquel appartient la théorie afin d'inférer, à partir du modèle étalon, des compléments théoriques.

3. L'examen de la correspondance avec les formes et les applications d'une théorie parente du domaine (habituellement antérieure, ayant une plus large extension mais comportant des faiblesses), afin d'inférer des développements théoriques complétant la théorie pour lui permettre de viser les applications couvertes par la théorie parente.

4. L'examen des présupposés, des implications ou des conséquences de la théorie et de ses opérationnalisations afin d'inférer les énoncés de présupposés, d'implications ou de conséquences qui seraient restés implicites et de dégager l'effet de cette explicitation en termes d'élargissement ou de restriction à l'étendue du domaine d'application de la théorie. Cette explicitation permet parfois aussi de connecter la théorie examinée avec d'autres énoncés théoriques et de renforcer la chaîne des énoncés.

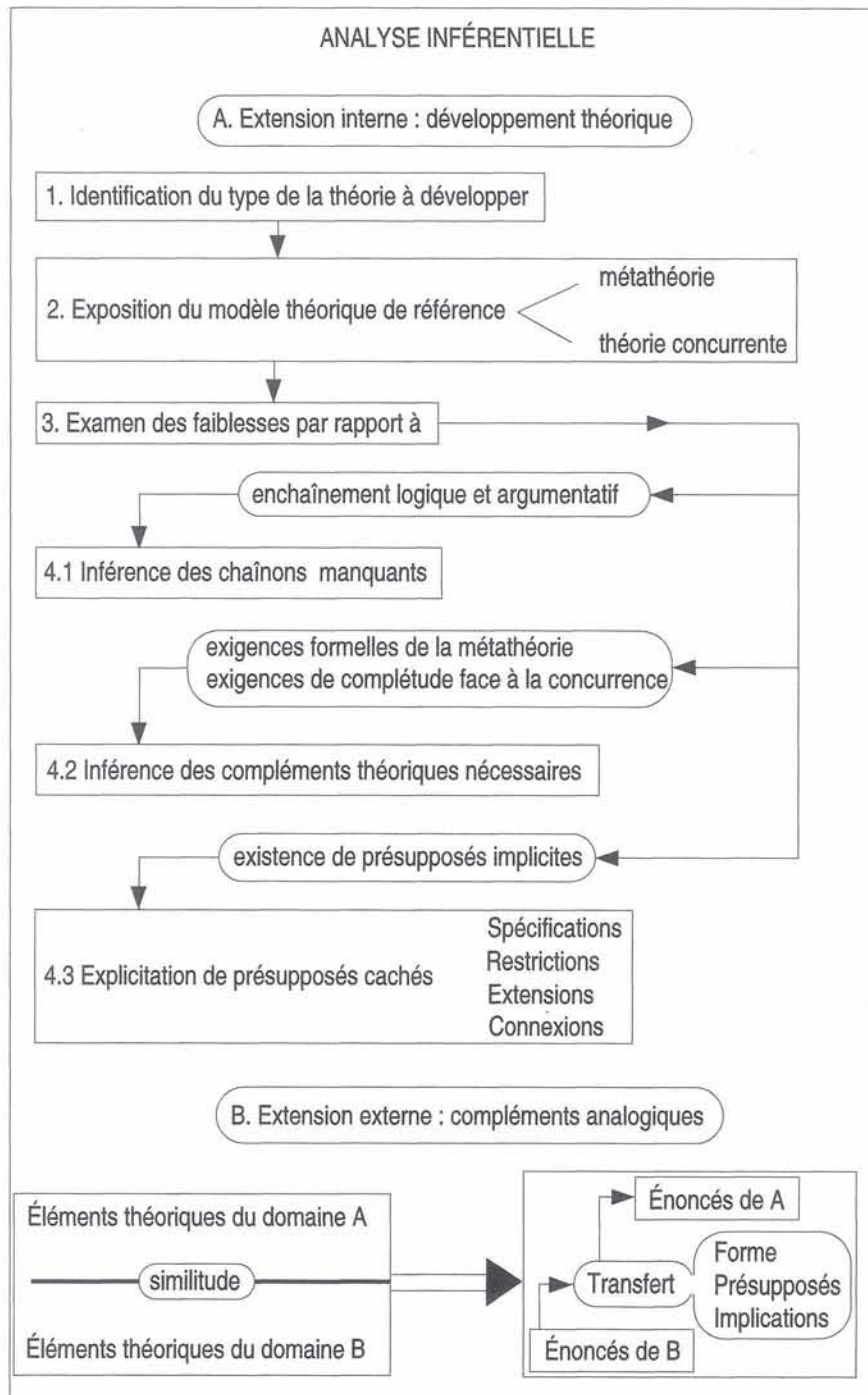


Figure 6.4 : L'analyse inférentielle.

Dans le cas d'une extension externe, les inférences procèdent surtout sur le mode analogique. La plupart du temps l'analyse inférentielle est effectuée par un chercheur du domaine A qui découvre les travaux d'autres chercheurs du domaine B et qui perçoit la possibilité de transférer à son domaine (A) une partie de leurs travaux, soit la méthodologie (ce qui implique toujours au moins quelques concepts théoriques), soit des résultats, c'est-à-dire quelques énoncés théoriques. L'analogie (c.-à-d. la similitude et la possibilité d'exprimer A par les termes de B) ayant été perçue par le chercheur à partir de ce qui s'énonce en B, le travail d'inférence doit d'abord se limiter à des formulations hypothétiques : si les énoncés valides en B s'appliquaient en A, quelle forme ces énoncés y prendraient-ils, quels devraient y être leurs présupposés, leurs implications et leurs conséquences. Une fois ce transfert hypothétique formulé, se pose la question de savoir quelle validité garderait la théorie B reformulée en A, étant donné les contraintes du domaine A et les transformations qui ont nécessairement été effectuées lors du transfert de B à A. L'épreuve de la nouvelle formulation de la théorie A doit alors suivre les mêmes voies que la mise à l'épreuve de toute nouvelle théorie, car ce n'est pas parce qu'elle est transférée de B où sa validité a pu être éprouvée, qu'elle bénéficie en A de la validité obtenue en B.

4.3 Les pièges de l'analyse inférentielle

Trois pièges se présentent dans l'analyse inférentielle par extension externe sur la base de l'analogie.

Le premier piège réside dans une mauvaise compréhension par le chercheur du domaine A des théories développées dans le domaine B. Étant donné que les concepts peuvent avoir un autre sens lorsque l'on passe du domaine A au domaine B, le chercheur du domaine A peut se méprendre sur leur sens dans le domaine B et percevoir une analogie là où il n'y en a pas. De plus, sans bien connaître le domaine B, le chercheur A peut ne pas percevoir tous les présupposés et toutes les restrictions liés aux énoncés du domaine B, et s'engager dans un transfert impossible. Aussi, pour éviter ce piège, le chercheur de A sera prudent de consulter un spécialiste de B afin de bien comprendre les exigences et le sens des énoncés développés en B.

Le deuxième piège s'observe lorsque des chercheurs du domaine B posent des inférences théoriques à partir de leur domaine (B) sur un autre domaine (A) qu'ils estiment semblable au leur, ou dont ils pensent qu'il pourrait n'être qu'un champ d'application du leur. Pour être plus clair, ce piège est celui où sont tombés quelques psychologues et médecins de grand

renom, lorsqu'ils ont pensé qu'à partir de leur savoir, ils pouvaient dire aux pédagogues ce qu'ils devraient faire. Or, ce piège est corrélatif au premier : le théoricien du domaine B ne connaît pas nécessairement toutes les contraintes et tous les présupposés qui régissent le domaine A. Il risque donc, sans consultation et sans coopération préalables avec les spécialistes du domaine A, de percevoir des analogies là où il n'y en a pas (parce que les contraintes diffèrent) ou de transférer des théories affaiblies lorsqu'elles prétendent s'appliquer en A. Les dégâts évidents produits par les importations brutales (elles allaient sauver l'éducation) venant de la psychologie, de la médecine, de la sociologie, de la linguistique et des mathématiques, devraient rappeler à la prudence chaque fois qu'une théorie externe à l'éducation apparaît séduisante.

Le troisième piège est semblable à celui évoqué pour l'analyse critique : le chercheur ayant effectué un transfert dont la puissance lui paraît évidente, doit s'assurer de la pertinence de ce transfert, le contraindre à l'examen de sa validité et de son efficacité, sans quoi d'autres chercheurs ne manqueront pas de prendre la théorie hâtivement transférée comme bouc émissaire de leurs railleries.

5 Les questions stratégiques

Insistons brièvement sur quelques problèmes observés pour que, prévenus, les chercheurs évitent de tomber dans des facilités qui transforment les écrits spéculatifs en écrits dogmatiques fort peu scientifiques.

5.1 Les fausses évidences de l'analyse du manque

L'absence d'une tradition qui oblige les chercheurs, les théoriciens et les producteurs de matériel pédagogique à déclarer leurs préconceptions (dont leur épistémologie, leur conception de l'éducation et les finalités et intentions de leurs travaux) conduit, entre autres, à des analyses critiques caricaturales et absurdes. En effet, prenant l'exemple d'un matériel pédagogique, il est probable que tout chercheur ne partageant pas la même conception et les mêmes finalités de l'éducation que le concepteur du produit, pourra facilement y trouver des manques et des excès, simplement en relevant les éléments opposés à ses préconceptions ou l'absence des éléments susceptibles de répondre à ses préconceptions. Si le modèle A implicite du concepteur ne correspond pas au modèle B implicite du chercheur, et si la non-correspondance des modèles A et B n'est pas déclarée et analysée, il est attendu que tout produit issu du modèle A aura des défauts s'il est évalué à partir du modèle B. Toute analyse ainsi faite est abusive,

mais elle sera jugée malhonnête si, de plus, elle fait comme si les préconceptions implicites sont partagées par les utilisateurs et par le concepteur. La rigueur intellectuelle impose, d'abord, de dévoiler les modèles implicites, de les comparer, puis d'examiner si les produits issus de chacun des modèles sont conformes aux modèles desquels ils sont issus. L'argumentation d'une évidence du manque ou de l'excès d'une théorie ou d'un produit ne peut se faire qu'à partir des préconceptions qui fondent le champ de cette théorie ou de ce produit. La discussion sur les préconceptions est vaine : ce sont des choix de valeurs qui ne se discutent pas, qui ne s'imposent pas, qui doivent être respectés.

5.2 De quelques stratégies de conviction

Toute recherche, spéculative ou autre, vise la production d'énoncés crédibles qui parviendront à convaincre les lecteurs de leur validité prétendument logique, bien qu'en fait plutôt psycho-sociologique. Pour atteindre à cette conviction, l'auteur d'énoncés peut recourir à divers arguments qui relèvent de stratégies amusantes et apparemment efficaces, mais dont l'efficacité n'est souvent qu'à court terme. Très utiles dans un débat oral pour faire taire un interlocuteur, ils résistent mal à l'analyse lorsque les auditeurs ont le temps de réfléchir ou le loisir d'analyser ces arguments, ce qui est toujours le cas des textes écrits. Nous examinons rapidement sept de ces stratégies.

5.2.1 L'argument étymologique

Utilisé dans une discussion où les présupposés et concepts n'ont pas été préalablement clarifiés, et qui semble s'embourber au désavantage d'un participant, l'argument étymologique coupe la discussion par le retour à une définition savante, apparemment indiscutable, à partir de laquelle toute la discussion devrait reprendre. Si le savant étymologiste est le seul dans le groupe à manifester cette maîtrise de la langue, l'argument emporte la victoire, les autres n'ayant pas la possibilité de reprendre toute la discussion en ajustant la signification et l'usage de leur terminologie. L'argument étymologique semble donc un argument terminal. Cependant, il est faible car, si l'étymologie peut bien indiquer la source d'un terme, elle ne garantit pas que son sens soit resté celui de ses origines; au contraire, elle permet de voir son évolution. Dès lors, si l'étymologie peut être un argument utile pour préciser le champ originare d'un concept, ce type d'argument devrait être présenté en début de discussion, pour savoir d'où on part, et non en cours de discussion pour la bloquer par un retour aux sources.

5.2.2 L'argument historique

Un peu similaire au précédent, l'argument historique a un impact sur les auditeurs-lecteurs étant donné le poids accordé au déterminisme historique. Étant donné l'histoire, il est évident que... La faiblesse de l'argument historique tient à ce que, si l'histoire témoigne souvent de l'évolution en continuité des événements, elle comporte aussi des ruptures, de révolutions, des changements sinon brutaux, du moins imprévus. Dès lors l'usage de l'argument historique ou celui de la tradition, souvent utilisé dans une perspective conservatrice, peut être retourné et même annulé par les ruptures qu'il porte aussi. Si l'un peut dire que «l'histoire nous montre que X» suivi d'un énoncé actuel, il ne sera pas difficile à un contradicteur de trouver quelques «faits historiques qui montrent à l'évidence que non X».

L'histoire est d'une telle richesse que les historiens trouvent toujours de quoi se contredire entre eux. L'appel à l'argument historique n'a souvent qu'un impact de courte durée, car il stimule souvent un contradicteur à entreprendre une recherche pour déterrer un fait ignoré qui renverse l'argumentation.

5.2.3 L'argument du bon sens ou du sens commun

Ce type d'argument ne devrait pas avoir sa place dans la recherche spéculative, d'autant plus que tout l'effort de la recherche scientifique consiste à démonter les préjugés du sens commun et les visières du bon sens. Le plus souvent, faire appel au bon sens ou au sens commun indique que l'on est à court d'argument valable pour soutenir un énoncé. Cet argument, populiste et démagogique, court-circuite la recherche et tente de faire appel aux émois majoritaires des non-spécialistes contre les arguments minoritaires des spécialistes. Cet argument polémique est de l'ordre de la politique.

5.2.4 L'argument anecdotique

C'est, dans la plupart des cas, l'argument le plus efficace. L'auteur capable de faire suivre un énoncé théorique d'une illustration anecdotique montrant comment la théorie fonctionne dans la banalité du quotidien, emportera l'adhésion de ceux que la théorie laissait froids. Bien des gens effectuent leurs choix non pas sur la base d'une évaluation rationnelle, mais par empathie avec ce qui a été vécu par une personne qu'ils connaissent bien et à laquelle ils accordent du crédit. L'auteur, comme l'enseignant,

capable de raconter une histoire comme si ce n'était pas une fiction mais le témoignage de la réalité, et de la rendre proche des émois de ses lecteurs-auditeurs, y fondera plus la crédibilité de la théorie que le rapport technique de vérifications statistiques ou expérimentales. Dans tous les secteurs, dont l'éducation, où les discours touchent souvent des objets susceptibles de réveiller des souvenirs et des émotions, l'argument anecdotique est un outil fort de conviction auprès des non-spécialistes. Aux yeux des spécialistes, l'anecdote n'a évidemment pas le même poids, car ils peuvent souvent montrer que l'anecdote produite peut tout aussi bien s'expliquer par une autre théorie, ou qu'elle n'est qu'un cas exceptionnel; dans une recherche spéculative l'argument anecdotique devra donc être soigneusement choisi et évalué avant d'être utilisé.

5.2.5 L'argument par le jeu de mots

Le jeu de mots est un argument qui fait brillant, qui donne l'impression (à ceux qui n'en sont pas victimes) d'une grande aisance et de la maîtrise du domaine en question. Sa faiblesse vient de ce qu'analysé à froid, il apparaît vite comme un simple jeu de langage dont on peut démonter les rouages et montrer les faiblesses logiques. Dans une conférence, il peut détendre, mais son abus dans un écrit fatigue. Utilisé avec économie, il peut amuser et convaincre, alors que répété, il produit, chez le lecteur qui analyse le texte, le sentiment que l'auteur est superficiel, que son texte manque d'assise, qu'il ne dispose pas assez d'arguments théoriques ou empiriques valides. En fin de compte, il choque le lecteur déçu de ne pas être pris au sérieux par l'auteur.

Étant donné ses faiblesses, le jeu de mot pourra faire partie d'une introduction, d'un paragraphe charnière où l'on se permet de faire un peu d'esprit pour marquer une transition et reposer un moment le lecteur, mais il ne devrait pas être utilisé dans le corps d'une argumentation. Là, le mot d'esprit tombe à plat.

5.2.6 L'argument de la logique et de la raison

L'appel à l'évidence logique comme l'appel à la raison sont des arguments qui semblent irrécusables. Qui, dans la recherche, voudrait être illogique et irrationnel ? Personne. Dès lors, des énoncés du genre «l'évidence logique nous conduit à...» ou «la raison nous montre que...» ont parfois un effet pervers : stimuler le lecteur à vérifier si l'auteur est logique, si ses conclusions ne sont pas hâtives et douteuses ou s'il ne dissimule pas quelques raisons qui rendent son raisonnement boiteux.

L'argument de la logique ou de la raison ne peut tenir que si une analyse de la validité logique des raisonnements a été effectuée au préalable. L'appel à l'évidence logique ou à la raison n'est efficace que lorsqu'il est l'aboutissement d'un travail d'analyse; il ne l'est pas du tout s'il sert de substitution à un tel travail.

5.2.7 L'argument de l'éthique et de la métaphysique

Ces arguments sont difficilement acceptables par les scientifiques teintés de positivisme, qui ne les considéreront que comme des moralisations. En fait, un peu à l'image de l'argument étymologique, les arguments éthiques et métaphysiques ont bien leur place dans les écrits, mais pas dans le corps de la recherche; ils doivent se situer au début du travail, lorsque le chercheur présente ses préconceptions, ou à la fin du travail, quand il en envisage des applications. Souvent même, on va exiger du chercheur qu'il soit capable de les mettre entre parenthèses, d'en faire provisoirement abstraction, afin que la théorisation puisse suivre son cours sans restriction aucune. On lui dira que c'est en fin de parcours, lorsque la théorisation est achevée, qu'il a le droit d'évaluer ses conclusions et leurs applications possibles à partir de ses options éthiques et métaphysiques.

Personnellement, l'histoire de la recherche scientifique en cette fin du XX^e siècle, nous conduit à penser que même si le chercheur a le droit d'évaluer le résultat de ses travaux en fin de parcours à partir de telles options, il a le devoir de les exposer avec ses préconceptions lors des explicitations préalables au début de ses travaux et des écrits qui les rapportent. Lors de la phase initiale, l'énoncé des arguments éthiques et métaphysiques, qui sont des arguments absolus, permettra au chercheur de mieux contrôler leur interférence avec la suite de son travail et de les voir comme des limites qu'il se donne et non pas comme des arguments qu'il impose aux autres lorsqu'il se sent dépassé par leur argumentation. Comme énoncés préalables, ils sont donc efficaces; intervenant en cours de travail, ils signalent une faiblesse donnant l'impression que le chercheur bat en retraite et qu'il se réfugie derrière des arguments d'un autre ordre (absolu) que ceux qui sont attendus.

LES DÉMARCHES DE LA RECHERCHE APPLIQUÉE

- 1 La recherche évaluative**
 - 1.1 L'évaluation critique*
 - 1.2 L'évaluation empirique*
 - 1.2.1 L'évaluation expérimentale
 - 1.2.2 L'évaluation par comparaison interne des paramètres et des conditions de la performance
 - 1.3 L'évaluation adaptative : amélioration et promotion*
- 2 La recherche-intervention**
 - 2.1 La recherche-action (enjeux politiques)*
 - 2.2 L'analyse de système et la méthode des systèmes souples (ou la recherche-action aux enjeux pragmatiques)*
 - 2.3 L'innovation et sa dissémination*
- 3 La recherche de développement**
 - 3.1 Le développement de concept*
 - 3.2 Le développement d'objet*
 - 3.3 Le perfectionnement des habiletés personnelles comme développement d'outils professionnels*

Les pratiques de la recherche appliquée sont guidées par quelques grandes intentions ou enjeux qui commandent des démarches spécifiques. On peut tracer une table à double entrée comportant les trois enjeux majeurs et les trois formes de la recherche appliquée. L'examen des méthodologies correspondant aux formes de recherche du tableau 7.1. se fera colonne par colonne. Rappelons que chacune de ces formes de recherche devrait comporter une phase de recherche spéculative grâce à laquelle le problème, même appliqué, pourra être mieux conceptualisé et mieux instrumenté.

Enjeux	Formes de recherche		
	Évaluation	Intervention	Développement
Politique Changement hétéro-centré	Décision	Recherche-Action : Révolution Émancipation-Appropriation Pastorale-Réforme	Développement de concept : application d'un concept dans la pratique + création d'un besoin
Pragmatique Solution fonctionnelle	Amélioration Adaptation	Analyse de système et résolution de problème : correction de dysfonctions	Développement d'objet : analyse de besoins et mise au point d'outils, de techniques, de stratégies
Ontogénique Changement auto-centré	Légitimation	Innovation Création	Développement personnel : Pratique-réflexive Heuristique Apprentissage dans l'action

Tableau 7.1 : Les enjeux et les formes de la recherche appliquée.

1 La recherche évaluative

La recherche évaluative a pour but de fournir des arguments de faits pour un jugement de valeur. Cette évaluation s'effectue en fonction de divers enjeux.

1° Politique : décider d'un choix, ce qui peut vouloir dire décider d'investir ou d'arrêter un investissement. Dans la plupart des cas, cette évaluation est sommative : elle fournit un portrait statique à partir de mesures de performance et d'une description du système.

2° Pragmatique : améliorer ou adapter une pratique ou un matériel. Dans cette perspective, l'évaluation est d'avantage formative : en plus des mesures d'états, elle examine la dynamique des processus et des actions.

3° Ontogénique : légitimer, renforcer une pratique déjà développée. Alors que, dans cette perspective, l'évaluation devrait être formative, elle se présente souvent comme sommative. Cependant, lorsque le chercheur ne tient pas à justifier sa pratique, l'évaluation devient souvent formative et se prolonge en recherche de développement.

Les évaluations politique et ontogénique peuvent suivre les mêmes démarches : une démarche spéculative, dite *critique* et une démarche empirique. Dans celle-ci, deux perspectives coexistent et parfois collaborent : un courant objectiviste privilégiant une méthodologie *formelle* et un courant constructiviste préférant une méthodologie plus *naturaliste* ou qualitative. La méthodologie générale de l'évaluation est la comparaison, et toute la question consiste à savoir qui compare quoi et avec quels critères¹.

1.1 L'évaluation critique

L'évaluation critique est utilisée autant par des décideurs (enjeu politique) que par des théoriciens qui veulent défendre une intuition théorique (enjeu nomothétique), ou par des promoteurs qui veulent justifier

1 Des nombreuses publications sur l'évaluation, on retiendra les ouvrages suivants pour une étude plus poussée de ce domaine :
Barbier J.M., *L'évaluation en formation*. Paris, PUF, 1985.
Guba E.G. et Y.S. Lincoln, *Fourth Generation Evaluation*. Sage Publication, CA : Newbury Park, 1989.
Lecointe M. et M. Rebinguet, *L'audit de l'établissement scolaire*. Paris : Les éditions d'organisation, 1990.
Lecomte R. et L. Rutman, *Introduction aux méthodes de la recherche évaluative*. Ottawa : Université de Carleton, 1982.
Nadeau M.A., *L'évaluation de programmes*. Sainte-Foy, QC : Les presses de l'Université Laval, 1988.
Paquette C., (éd.), *Des pratiques évaluatives*. Victoriaville, QC : Editions NHP, 1984.
Patton M.Q., *Qualitative Evaluation and Research Methods*. Deuxième édition. Sage Publication, CA : Newbury Park, 1990.
Shadish W.R., Cook T.D. et L.C. Leviton, *Foundations of Program Evaluation, Theories of Practice*. Sage Publication, CA : Newbury Park, 1991.

leur option (enjeu ontogénique). Elle consiste le plus souvent à comparer le contenu et la présentation d'un matériel pédagogique (méthodes, programme, stratégie, outil, etc.) avec un idéal. Celui-ci est un modèle plus ou moins déclaré de ce que devrait être ce matériel ou tout matériel pédagogique. L'évaluation critique procède selon les stratégies de l'analyse spéculative appliquée sur le matériel à critiquer en prenant le modèle idéal comme étalon.

Dans ce genre de critique, le modèle étalon réfère le plus souvent à une conception et à des valeurs que l'auteur du matériel ne partage pas avec le critique. La condamnation est alors facile : le critique n'a même pas à examiner la cohérence ou la qualité intrinsèque du matériel, il lui suffit de relever les manques, les lacunes ou les excès par rapport au modèle. Mais en général, l'écart entre le modèle du critique et le modèle implicite au matériel est si vite évident, que le lecteur peut avancer une hypothèse sur les conclusions avant d'avoir terminé la lecture. L'utilité de l'évaluation critique dépend du degré de communauté des valeurs et des conceptions partagées par l'auteur, le critique et les utilisateurs potentiels : si ces derniers adhèrent aux mêmes valeurs que le critique, ils lui accorderont foi; s'ils ont les mêmes conceptions que l'auteur du matériel, ils rejeteront la critique. De toute façon, un matériel pédagogique est légitime de par la manière dont il exprime et réalise les valeurs, les idéaux, les normes du groupe social dans lequel l'école est insérée, et l'examen de cette adéquation ne prend pas une recherche évaluative.

1.2 L'évaluation empirique

L'évaluation empirique constitue le plus souvent un projet émanant de praticiens ou d'administrateurs qui cherchent à savoir s'ils ne se sont pas trompés dans la construction ou la mise en place d'un *matériel* (nous utiliserons souvent ce terme pour les divers objets d'une évaluation : système, programme, curriculum, manuel, stratégie, outil didactique). Dans le courant objectiviste, plusieurs stratégies ont été élaborées : la comparaison expérimentale et diverses procédures de comparaison interne.

1.2.1 L'évaluation expérimentale

La stratégie spontanément privilégiée par les non-spécialistes de l'évaluation qui veulent démontrer ou prouver la valeur d'une intuition pédagogique ou d'un matériel consiste à monter une expérimentation dans laquelle on compare les résultats obtenus grâce au matériel choisi avec les

résultats obtenus avec un matériel classique. Deux cas se présentent le plus souvent, soit l'expérimentation avec groupe contrôle (témoin) et groupe expérimental, ou l'expérimentation par classe unique avec retrait du matériel (traitement) (voir chapitre 9 : les plans de recherche). Ces deux plans expérimentaux d'évaluation, inspirés de la psychologie, ne sont pas sans poser des problèmes qui les invalident.

En ce qui concerne la procédure avec groupe expérimental et groupe témoin², on peut résumer les problèmes en posant les questions suivantes :

- Les objectifs des deux matériels et les groupes d'élèves sont-ils équivalents ? Les professeurs ont-ils la même expérience, les mêmes compétences, la même motivation ? Comme il est difficile sinon impossible d'obtenir des groupes constitués au hasard, il est quasi impossible d'assurer la comparabilité initiale des groupes.
- L'expérimentation n'est-elle pas contaminée par les attentes mises dans le nouveau matériel, par son aspect de nouveauté, ou par la crainte que l'évaluation du matériel ne serve en même temps à juger le personnel ? Ces variables psychosociales (voir chapitre 10 : les biais psychosociaux) biaisent suffisamment l'expérience pour qu'on ne sache plus, en fin de compte, si les résultats observés sont attribuables au nouveau matériel ou aux attentes qu'il suscite, à la nouveauté, à la crainte, etc.
- Les facteurs (ou variables) parasites de la validité interne ont-ils pu être contrôlés (voir chapitre 9) : l'histoire, la maturation, la sélection, l'apprentissage du test, la réactivité à la situation expérimentale, la constance des consignes et des mesures, la perte de sujets qui lâchent, les effets de régression statistique, etc. ? Comme les réponses à ces questions sont souvent incertaines, les conclusions le sont aussi.
- Les sujets et les conditions de l'expérience sont-ils représentatifs et peut-on étendre les conclusions à d'autres contextes ? Dans la majorité des cas, les particularités de la situation expérimentale ne coïncident que très partiellement avec les contraintes concrètes de l'utilisation du matériel, si bien qu'il devient difficile de préciser les conditions et les limites du transfert des conclusions de l'expérience vers le terrain scolaire ordinaire.

2 Van der Maren J.M., Des faiblesses du modèle expérimental à l'évaluation adaptative-interactive des interventions en éducation. *Revue canadienne de l'éducation*, 1986, Vol.11, n° 1, p. 72-82.

Les plans d'expérience ne comportant qu'un groupe unique auquel on applique et on retire le matériel à évaluer selon une séquence d'alternances, posent aussi des problèmes, dont les suivants :

- La performance ou la non-performance peut se maintenir en dépit du programme parce que le sujet a une mémoire, il se développe sur le plan cognitif, il a des intentions et une autonomie insoupçonnée;
- L'alternance du traitement et du retrait pose une question éthique : le retrait du matériel ne provoque-t-il pas quelques difficultés aux sujets ? Si le matériel est bénéfique pour les sujets, de quel droit peut-on les en priver, ne serait-ce que momentanément ? A-t-on l'accord des sujets pour une telle procédure ?
- Les modifications observables seront-elles dues à l'apprentissage, au développement, à la présentation du matériel ? Dans une situation complexe avec un sujet intentionnel, l'utilisation d'un plan à cas unique permet difficilement d'isoler l'effet du matériel, la mémoire du sujet, la rapidité de ses réactions au changement et, en conséquence, de savoir exactement ce qui a été évalué.

Enfin, la mesure impliquée par ces plans expérimentaux pose aussi les problèmes qui ont été soulevés lorsque l'on a traité des paradoxes (chapitre 4), entre autres ceux du hasard, de la quantification avec le dilemme de la pertinence des pré- et post-tests, et du temps. Dès lors, pourquoi se donner tant de mal à monter une recherche évaluative par expérimentation lorsque l'on sait que, de toute façon, on ne pourra pas démontrer la valeur du matériel et que l'auteur du matériel restera convaincu (et convaincant) d'avoir produit le meilleur ! En outre, aucune recherche ne justifie ni ne légitime un matériel pédagogique ; les jugements de valeurs, les jugements moraux et esthétique, de même que l'utilité d'un matériel ne peuvent être tranchés par une expérimentation scientifique.

1.2.2 L'évaluation par comparaison interne des paramètres et des conditions de la performance

Dans les stratégies d'évaluation développées par les spécialistes, la comparaison est avant tout interne, car ils savent que leurs travaux ne pourront rien prouver et qu'ils pourront seulement apporter des informations à partir desquelles un jugement pourra être porté. Dès lors, la stratégie à la base de l'évaluation est la confrontation des prétentions d'un matériel y compris les conditions qu'il exige, avec ses réalisations, étant donné les conditions qui lui sont accordées. Avant d'évaluer un ou plusieurs matériels, il convient que l'évaluateur précise l'usage attendu du matériel à choisir : il

s'agit d'établir un *cahier des charges* avant d'examiner les produits proposés. Dès lors, si pour un matériel donné (outils, programme, manuel, etc.) on veut comparer ce qui est obtenu ou les réalisations, avec l'attendu ou les prétentions, il faut préciser les exigences : les types, les critères et les conditions ou les contraintes de la performance ainsi que les effets secondaires.

Trois *types* de mesure de la performance doivent être distingués pour choisir les mesures les plus pertinentes dans le contexte de l'utilisation du matériel à choisir. Dans certains cas, les trois types de mesure devront être envisagés :

- *L'effectivité* se définit comme la proportion d'objectifs atteints (aspect quantitatif) ;
- *L'efficacité* se rapporte au niveau, à la manière ou à la qualité de l'atteinte (aspect qualitatif) ;
- *L'efficacités* est une mesure du rendement, c.-à-d. du rapport des coûts ou des investissements (financiers, sociaux, psychologiques) sur les bénéfices ou sur les gains.

Les *critères de performance*, c.-à-d. la détermination des objectifs et des actions (ou des productions) considérés comme indices de la réalisation des objectifs, peuvent provenir de deux sources.

- Les critères sont extrinsèques lorsque les prétentions ou les attentes à l'égard du matériel sont énoncées en dehors de sa construction. Par exemple, lorsque le matériel doit atteindre ou prétend atteindre les objectifs qui ont été décidés par le ministère ou les commissions scolaires sous forme de programme-cadre.
- Les critères sont intrinsèques lorsque c'est le contenu et la construction elle-même du matériel qui contient un énoncé de ses prétentions (objectifs et actions qui en témoignent).

Les *conditions de la performance* à inscrire au cahier des charges sont nombreuses et demandent souvent que l'on procède à une modélisation du système ou du matériel à évaluer. Cette modélisation comportera entre autres :

- Les acteurs ou les utilisateurs avec leurs caractéristiques, leurs compétences, et leurs attitudes (attrait, indifférence, résistance);
- L'organisation impliquée sur le plan matériel (modifications exigées aux locaux, matériaux, etc.) et sur le plan du travail (habitudes, conventions collectives, etc.);

- Le support de l'environnement humain interne à l'école (administration, cadres, personnel parascolaire et de soutien) et externe à l'école (parents, politiciens, commissaires, églises, syndicats).

En plus de la performance attendue, le cahier des charges doit aussi préciser quels sont *effets secondaires* tolérables ou admissibles et intolérables ou inadmissibles, à court, à moyen et à long termes, que ces effets soient prévisibles ou imprévisibles. Cela implique un système inscripteur ouvert à l'observation des effets autres que ceux qui sont attendus; ce qui veut dire que l'on doit, à l'avance, préciser les effets secondaires que l'on ne veut pas voir apparaître et l'instrumentation qui permettra de les détecter.

Le jugement évaluatif résultera ensuite de la comparaison des exigences du cahier des charges préalablement établi avec les données obtenues à partir des réalisations. Lorsque l'évaluation est préalable à un choix entre plusieurs systèmes ou matériels, la recherche évaluative comprendra, pour chacun, une comparaison interne sur

- ses objectifs, soit ce qu'il prétend faire;
- sa performance et ses effets secondaires, soit ce qu'il fait effectivement, et comment il le fait;
- ses conditions, soit ce qu'il exige pour pouvoir le faire.

Il faudra veiller à la comparabilité des cas, autrement dit à la standardisation des critères utilisés pour l'ensemble des comparaisons internes.

L'évaluation par comparaison interne privilégie une collecte de données la plus objective possible, car la mesure ou l'évaluation ne doit pas être suspectée de préjugés. Cependant, les évaluateurs considèrent de plus en plus qu'à côté des données obtenues par des techniques instrumentales (mesures objectives), ils doivent tenir compte d'informations subjectives obtenues directement du terrain. En effet, pour comprendre comment un matériel peut être plus ou moins performant, il faut souvent avoir accès aux représentations que les utilisateurs ou les acteurs ont des objectifs poursuivis, du fonctionnement et des conditions du matériel, car ces représentations peuvent différer de la conception qui a présidé à la construction du matériel. Si la collecte d'informations subjectives autant que de données objectives, permet aux décideurs de mieux identifier les raisons d'un choix, elle permet aussi de mieux expliquer aux acteurs les raisons d'un choix dont ils se sentiront moins victimes. Sans devenir pour

autant formative, l'évaluation justificative comporte donc de plus en plus de collecte de données par observation participante et par entretiens; elle devient de la sorte plus *participative*.

1.3 L'évaluation adaptative : amélioration et promotion

L'évaluation pour fin d'*amélioration* s'effectue selon une procédure d'*évaluation adaptative-interactive*. Il ne s'agit pas de comparer à d'autres, mais bien de partir de ce dont on dispose et d'examiner comment on pourrait le corriger, l'adapter afin d'atteindre ce que l'on souhaite. Tel que schématisé dans la figure 7.1, on procède par des boucles de comparaisons internes et de modifications successives permettant des ajustements progressifs dans une perspective *cybernético-négociative*. En effet, si l'on veut réellement obtenir une amélioration, chaque correction envisagée doit faire l'objet d'une négociation avec les acteurs qui devront la vivre; la correction imposée sans négociation est trop souvent boycottée par les acteurs pour que l'on puisse éviter de la négocier. De plus, c'est souvent par une discussion avec les acteurs que l'on pourra mettre en place des modalités efficaces d'adaptation du matériel. Autrement dit, l'évaluation adaptative ne peut se réaliser qu'en interaction avec les acteurs et les utilisateurs : elle doit être participative. L'évaluateur ne peut pas se contenter d'établir le diagnostic et de prescrire le remède, il doit les élaborer avec la participation des acteurs et participer à l'application du traitement. C'est seulement en étant impliqué sur le terrain que les boucles évaluatives-adaptatives pourront être efficaces.

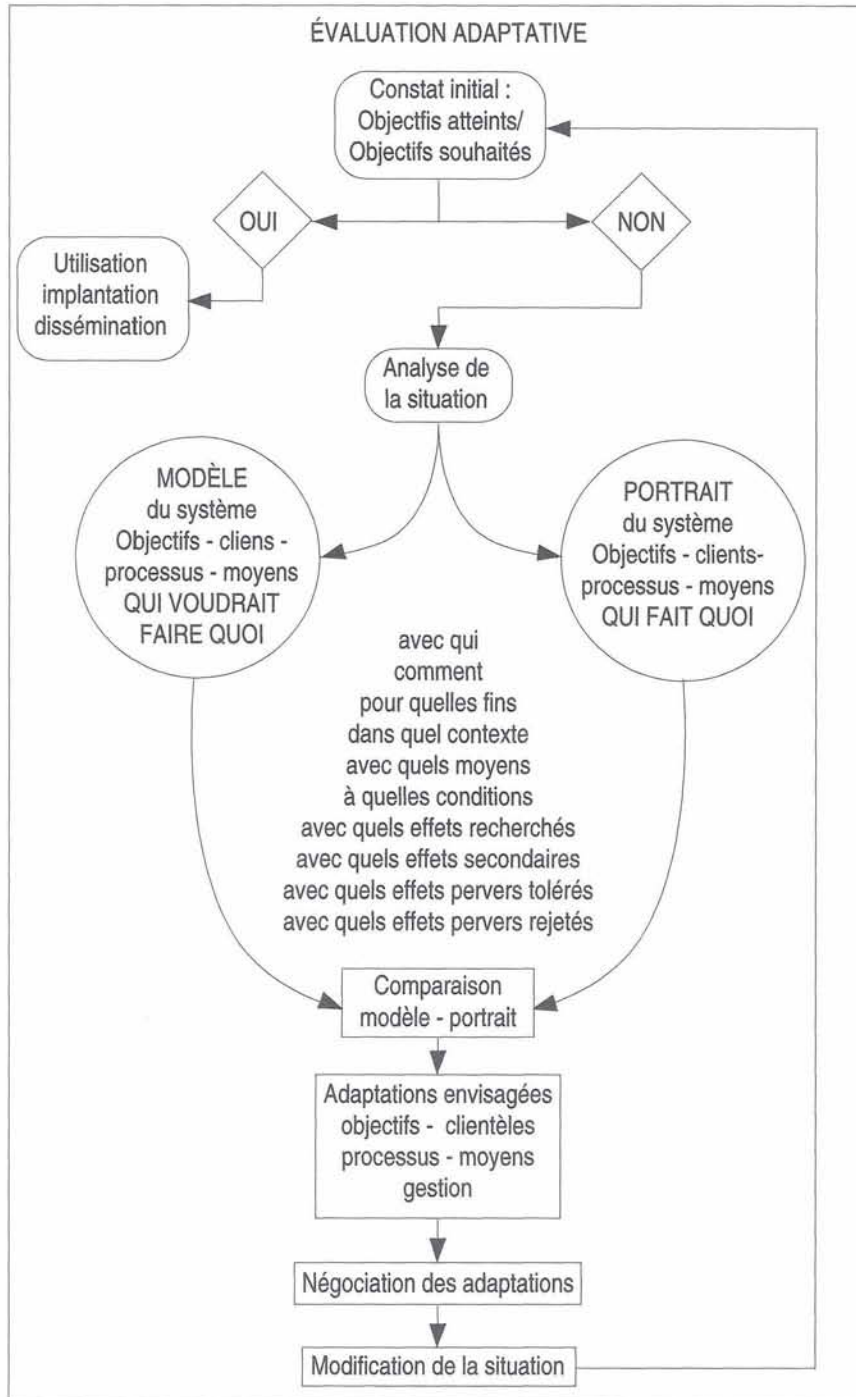


Figure 7.1 : Les étapes de l'évaluation adaptative.

2 La recherche-intervention

2.1 La recherche-action (enjeux politiques)

La recherche-action, que nous distinguerons de l'analyse de système qui en est une version aux enjeux pragmatiques (analyse fonctionnelle de système et résolution de problème), vise le changement plus ou moins radical d'une situation d'éducation avec la participation relative des populations concernées et des intervenants impliqués. Son principe reçoit des colorations différentes selon l'idéologie des auteurs et des acteurs. Pour les marxistes on peut en résumer les étapes par la trilogie «enquête-théorie-pratique», et à peine en caricaturer la stratégie en la ramenant à l'exacerbation des contradictions. À la suite des travaux de Lewin, les psychologues procèdent d'abord à l'analyse de la situation pour identifier les individus, les moments et les occasions ou situations susceptibles de changement (ce qu'en leur jargon, ils appellent : identifier les portes et les portiers). Ensuite, leurs interventions visent à déstabiliser les normes des groupes et les habitudes des individus pour, enfin, réinstaurer de nouvelles normes qui seront consolidées par une organisation pratique efficace.

Depuis les premiers essais documentés par Lewin, diverses formes et idéologies ont abouti à une assez grande variété de pratiques en recherche-action, ce qui s'explique d'autant plus que chaque recherche particulière, si elle se veut efficace, doit tenir compte des particularités du contexte de la demande et de son déroulement. On rencontre des pratiques qui se situent aux extrémités de l'axe idéologique, allant de pratiques humanistes³ jusqu'aux pratiques plus critiques et dialectiques⁴, en passant par un certain fonctionnalisme⁵. Goyette et Lessard⁶ ont produit une analyse des pratiques québécoises en recherche-action qui permet de situer les pratiques, les enjeux

- 3 Morin A., Réflexions sur la recherche-action à partir d'une expérience d'animation. *Prospectives*, XX, 1-2, 1984, p. 49-54. Voir aussi Gagnon R., Typologie et stratégies de recherche-action. *Prospectives*, XX, 1-2, 1984, p. 42-48. Morin A., *Recherche-action intégrale et participation coopérative*. vol. 1 : Méthodologie et études de cas; vol. 2 : Théorie et rédaction du rapport. Montréal, QC : Éditions Agence d'Arc, 1992.
- 4 Horth R., *L'approche qualitative comme méthodologie de recherche en sciences de l'éducation*. Pointe-au-Père, QC : Les éditions de la MER, 1986.
- 5 Gélinas A. et Ch. Gagnon, Systémique, recherche-action et méthodologie des systèmes souples. Chicoutimi, QC : GRIR/UQAC, vol. II, n° 1, mars 1983. Pour une formation à cette démarche, voir : Claux, R. et A. Gélinas, *Systémique et résolution de problèmes, selon la méthode des systèmes souples*. Montréal, QC : L'Agence d'Arc, 1982.
- 6 Goyette G. et M. Lessard-Hébert, *La recherche-action, ses fonctions, ses fondements et son instrumentation*. Sillery, QC : PUQ, 1987.

et les références idéologiques de ce type de recherche. Actuellement⁷, les pratiques de la recherche-action ont tendance à réduire l'aspect politique du projet pour devenir des outils d'intervention andragogiques et psychopédagogiques. Par ailleurs, étant donné l'insertion de l'action dans ce type de recherche, son évaluation et sa validation posent quelques problèmes à propos desquels F. Gauthier et C. Baribeau⁸ ont proposé des pistes de réflexions et de pratiques.

Les principales phases d'une stratégie de changement sont schématisées à la figure 7.2; elles doivent être respectées quelle que soit l'idéologie qui justifie le changement. Cette stratégie s'effectue le plus souvent en travaillant avec un groupe d'individus dont on sait que les normes ne sont pas homogènes. En fait, la recherche-action consiste justement à affecter la dynamique des rapports de cohésion et des rapports de force de façon à modifier les normes de la majorité des membres.

Le modèle de la recherche-action schématisé à la figure 7.2 comporte quatre grandes tâches, suivies chacune d'une ou deux étapes qui doivent être quasi complétées avant d'entreprendre la tâche suivante. Ainsi, lorsqu'on veut changer les normes d'un groupe, il faut au préalable avoir très bien délimité les changements que l'on souhaite provoquer (étape 1). Plus d'une recherche-action avorte parce que ses objectifs sont flous, mal définis ou qu'ils n'obtiennent pas l'adhésion de tous les chercheurs impliqués, d'autres n'aboutissent pas parce que les marginaux ne sont pas assez supportés (étape 3c) ou parce que ceux qui ont voulu le changement, ne font pas tout pour mettre en place une organisation pratique suffisamment efficace pour renforcer les nouvelles normes (étape 4).

Après l'identification précise des changements visés, la deuxième phase consiste à analyser minutieusement les comportements, les interactions et les rôles qui se développent dans le groupe en tenant compte des différentes tâches qu'il tente de réaliser. Dans tous les groupes, il y a des activités et des moments où certains membres sont plus ou moins d'accord avec les décisions du groupe : même dans une bande, il y a des marginaux. La première tactique consiste donc à identifier quels sont les marginaux du groupe, à quels moments et pour quel genre de tâches ils sont susceptibles de manifester leur marginalité, c'est-à-dire de constituer une minorité

7 Baribeau C. (éd.) La recherche-action de Kurt Lewin aux pratiques contemporaines. Numéro spécial de la *Revue de l'Association pour la recherche qualitative*, Volume 7, automne 1992.

8 Gauthier F. et C. Baribeau, Traitement de la qualité d'un plan de recherche-action. Dans Cl. Paquette (éd.) *Des pratiques évaluatives*. Victoriaville, QC : Editions, NHP, 1984, p. 285-322.

dominée par rapport au reste du groupe. On postule aussi que tous les groupes sont guidés par une minorité dominante (les chefs, les leaders).

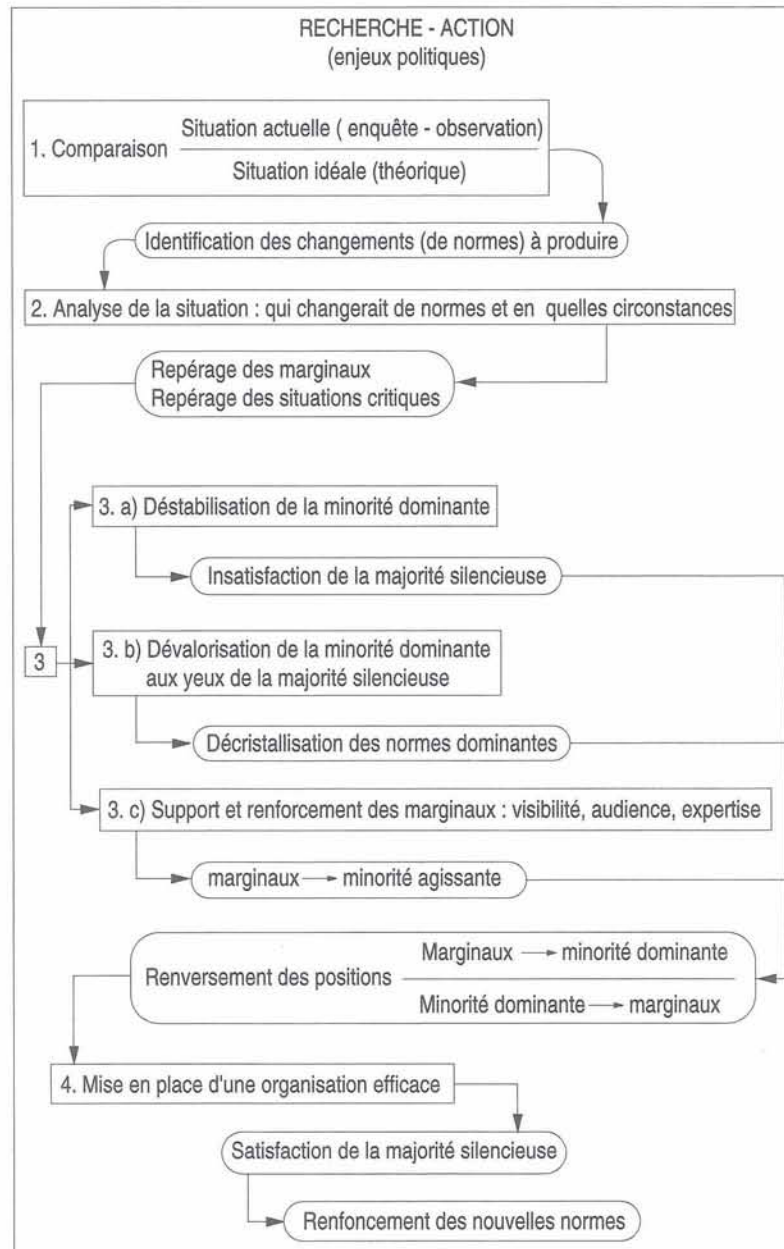


Figure 7.2 : Les étapes de la recherche-action.

Parallèlement au repérage des marginaux (la minorité dominée), il est important de tenter de déstabiliser la minorité dominante face à la majorité : cela peut se faire soit en lui rendant plus difficile de satisfaire les membres par des obstacles que l'on pose à la réalisation de ses activités, soit en amplifiant l'impact et la signification des échecs que le groupe peut rencontrer. Si le groupe apparaît très homogène, dominé par des personnes qui ont un grand pouvoir et qui oppressent les autres, il faudra sans doute provoquer des évaluations de la performance des leaders par des contestations et par des questions parfois insidieuses. Par exemple, on peut demander pourquoi telle décision a été prise, et émettre l'hypothèse que si une autre décision avait été prise, le gain pour l'ensemble du groupe aurait été meilleur, mais que c'est par intérêt personnel ou par manque de compétence que les leaders ont choisi la décision qu'ils ont prise. Avec la déstabilisation des leaders, il importe aussi de mettre en doute les règles ou les principes selon lesquels ces leaders agissent : cela permet de décristalliser les normes dominantes, de les rendre plus fragiles, de permettre que soit rompue une loi du silence. Après avoir laissé passer quelques doutes du genre, il n'est pas rare de voir un dominé oser poser une question qui peut susciter l'apparition d'autres marginaux : une minorité dominée se réveille et va devenir agissante. Autrement dit, tout en repérant ceux qui constituent une minorité sur laquelle on pourrait s'appuyer lors d'un changement, on tente de provoquer un mécontentement, une insatisfaction de la majorité vis-à-vis de la minorité dominante.

Lorsque les marginaux se révèlent une minorité, il faut immédiatement les renforcer et leur offrir du support afin qu'ils constituent une minorité agissante. Pour ce faire, on veillera

- à leur donner l'occasion de se faire souvent entendre;
- à les appuyer pour qu'ils aient une visibilité telle que la majorité ne puisse les réduire physiquement ou moralement au silence;
- à leur fournir la documentation, les informations et les expertises nécessaires pour que, dans leurs discours et par leurs actions, ils puissent présenter des arguments convaincants qui vont miner la crédibilité de la minorité dominante et qui vont grossir leurs propres rangs.

L'expertise fournie à la minorité agissante servira aussi de cadre de référence aux normes qui se mettront en place à la suite du changement, afin qu'elles correspondent à celles que l'on souhaite voir s'établir. Elle sert aussi à la formation de la minorité agissante de telle sorte que lors du changement, les nouvelles décisions et les nouveaux leaders soient

rapidement efficaces, c'est-à-dire qu'ils obtiennent rapidement des succès qui satisfassent les membres et renforcent leur leadership.

Une fois ces tactiques enclenchées, le mouvement va progresser, pourvu que se maintienne le discrédit de la minorité dominante aux yeux de la majorité. On continuera donc de provoquer des difficultés, d'amplifier l'insatisfaction et de soutenir la minorité agissante par des moyens matériels et des appuis qui renforcent sa crédibilité. Arrive le moment du changement : la crédibilité et les rangs de la minorité agissante se sont gonflés au point où celle-ci se sent assez forte pour acculer l'ancienne minorité dominante à abandonner le leadership du groupe ou à modifier ses normes. Cette transition sera plus ou moins violente selon l'attachement des anciens leaders à leur pouvoir sur le groupe, selon la force de l'oppression qu'ils exerçaient et selon le degré de rationalité ou d'émotivité des discussions.

Dès que le changement se produit, il convient d'agir de façon à ce que les nouveaux leaders puissent mettre en place avec succès les normes souhaitées, c'est-à-dire en réalisant des actions qui satisfont la majorité des membres, et de façon à ce que les anciens leaders puissent se rallier aux nouvelles normes que le groupe se donne. Pour réaliser cela, il serait bon que les nouveaux leaders, appuyés par la majorité renouvelée et par les chercheurs-intervenants, distinguent entre la personne et la conduite des anciens leaders, qu'ils soient assez subtils pour utiliser les compétences de ces derniers et pour, très rapidement, amener tout le groupe à des progrès qui vont lui donner une nouvelle identité, une nouvelle cohésion. Cela implique que l'on ne fasse pas des anciens leaders les nouveaux boucs émissaires du groupe.

Dans les cas où les anciens leaders se maintiennent en poste, se réforment et acceptent que les normes soient changées, il faudra tenir éveillée la vigilance de la minorité qui a provoqué le changement et renforcer le changement par des succès, afin de stabiliser les nouvelles normes obtenues.

La stratégie de changement dont on vient de brosser le tableau est d'un emploi délicat en situation scolaire. Appliquée à une classe qui pose des problèmes, elle peut échouer et, dès lors, renforcer la déviance de la classe et être source de perturbations dans toute une école. Même lorsqu'elle est utilisée avec des groupes d'adultes, dans ou hors d'une institution ou sur tout un système scolaire, il importe que le chercheur qui intervient mesure la pertinence écologique des changements qu'il introduit, sinon, une fois l'intervention terminée, le groupe sera encore plus démuni face à son environnement et aura le sentiment d'avoir été manipulé par le chercheur.

De toute façon, son utilisation requiert la collaboration d'un tiers expert dont les fonctions sont l'analyse et le reflet de ce qui se passe, et le contrôle de la tentation de manipuler le groupe ou le système pour des fins personnelles.

On peut facilement imaginer qu'une variante individualisée de la recherche-action est applicable lorsqu'il s'agit de modifier les représentations et les actions d'une personne. L'intervention devient alors plus psychopédagogique ou andragogique. Dans ce cas, il s'agira d'identifier quelles sont les activités de la personne dans lesquelles ses représentations sont les moins efficaces, de les déstabiliser en mettant en évidence leurs difficultés, d'amener la personne à essayer d'autres solutions efficaces, etc.

Telle que décrite ci-dessus, la recherche-action semble ne comporter que des actions, à part la phase d'analyse de la situation. Pour qu'une telle entreprise soit le lieu d'un apprentissage ou d'une construction de connaissance, il importe que les chercheurs et les acteurs gardent une trace analysable de ce qui s'est produit pendant toute la recherche. Il n'y a pas de recherche dans la recherche-action sans une chronique de l'action. Cette chronique, qu'elle soit sous la forme de mémos ou de journaux de bord, s'attachera à rapporter, au moment de l'action, toutes les modifications apportées au plan de l'action avec les conditions qui les ont provoquées et les raisons (analyses, réflexions et décisions) qui ont motivé les modifications. Les raisons écrites doivent impérativement être celles qui ont présidé à l'action; si la modification était de l'ordre d'une adaptation spontanée, le chercheur se gardera de lui attribuer une raison après coup. Comme la conduite d'une recherche-action est souvent plus sinueuse que rectiligne, la transparence de la recherche impose aux chercheurs de noter soigneusement le cheminement des perceptions et des analyses successives de la situation. Mais dans ce rapport, les perceptions, les analyses, les décisions doivent être notées comme des faits qui se succèdent et non comme des relations causales. L'hypothèse de relations causales découlera de l'analyse réflexive de la chronique qui se fera une fois la recherche-action terminée.

2.2 *L'analyse de système et la méthode des systèmes souples (ou la recherche-action aux enjeux pragmatiques)*

Divers courants, inspirés par l'analyse des systèmes⁹, proposent des démarches assez proches les unes des autres. D'une manière générale, elles constituent des compromis entre la démarche de résolution de problèmes,

9 Checkland. P., *Systems Thinking, Systems Practice*. New-York : Wiley, 1981.

l'ingénierie des systèmes-machines et l'analyse des systèmes humains. Nous décrivons plus explicitement la démarche que nous avons voulu respecter dans la direction d'une recherche-intervention en milieu collégial¹⁰.

Telles que représentées dans l'organigramme de la figure 7.3, les étapes en sont les suivantes :

1. Analyse de la demande (ce qui ne va pas) et des objectifs du système demandeur (ce qui se passe, ce qu'ils veulent, ce dont ils disposent, ce qu'ils peuvent faire) et des objectifs de l'intervention (ce qu'ils attendent et ce pour quoi le chercheur accepte de tenter une recherche de solution);

2.a. Conceptualisation de l'organisme demandeur en tant que système;

2.b. Modélisation du système et de son fonctionnement, c'est-à-dire en construire une représentation générale (voir chapitre 11 : Modélisation et simulation) afin de pouvoir effectuer des simulations;

2.c. Simulation (concrète, formelle ou imaginaire) du système et de son fonctionnement, c'est-à-dire attribuer une valeur particulière à chaque élément du modèle (c.-à-d. en construire un portrait hypothétique, s'en faire une image) pour estimer quel serait le résultat obtenu si le système fonctionnait avec de telles valeurs. Ensuite, comparer différentes simulations du fonctionnement du système (en attribuant d'autres valeurs à ses éléments) afin de choisir les valeurs optimales à accorder à chaque élément. Cette comparaison de simulations permettra d'élaborer les modifications à envisager. Même lorsqu'il s'agit d'une simulation imaginaire, il faut limiter l'ampleur, la fréquence et la direction des variations simulées de chaque élément du modèle aux limites des variations possibles dans le contexte (environnement) réel du système demandeur;

Dès cette étape et pour chacune des suivantes, il est essentiel que les chercheurs tiennent une chronique de la recherche. Cette chronique doit être enregistrée et fixée le plus vite possible après le déroulement de l'action, elle doit respecter au mieux les formes de l'action et n'avoir que les apparences d'un scénario : on y raconte comment les choses se déroulent, dans leur simple succession et dans leur décor, et sans inférer, alors qu'elles n'y étaient pas, des relations causales entre les événements. La chronique ne peut pas être la construction d'une histoire dans laquelle les actions s'enchaînent grâce à leurs motifs; elle est seulement le rapport de la succession des actions et des décisions;

¹⁰ Gattuso L., Les mathophobes, une expérience de réinsertion au niveau collégial. Mémoire de maîtrise. Faculté des sciences de l'éducation, Université de Montréal, 1987.

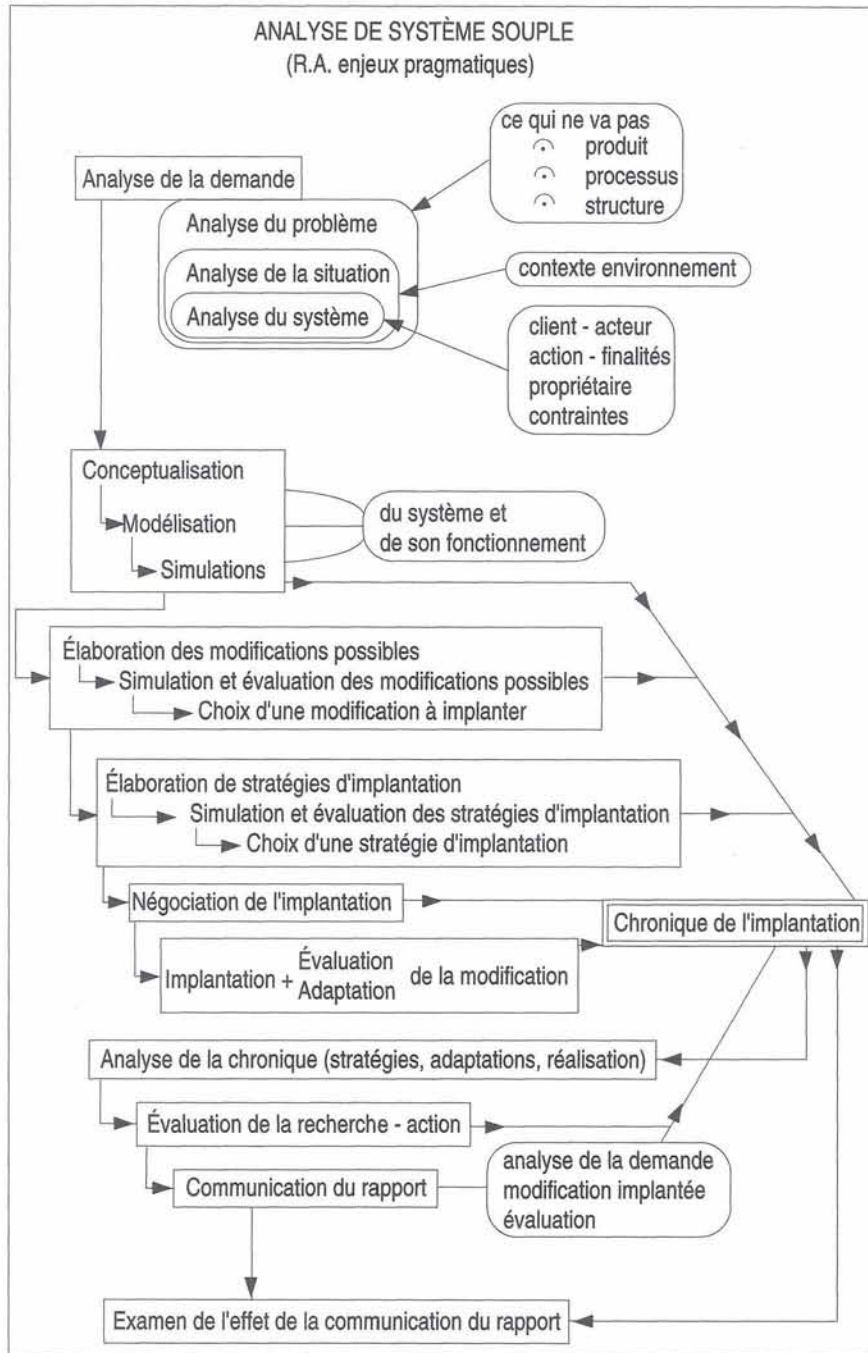


Figure 7.3 : Les étapes d'une analyse de système souple.

3. Élaboration des modifications possibles, simulation et évaluation des effets des modifications possibles, choix d'une modification à implanter. Dans une perspective de résolution de problème, il importe que plusieurs modifications soient envisagées et évaluées plutôt que de s'engager dans la réalisation de la première intuition;

4. Élaboration, évaluation et choix des stratégies d'implantation de la modification à implanter. À nouveau, plusieurs stratégies doivent être envisagées et évaluées, la plus évidente n'étant pas nécessairement la meilleure lorsqu'il s'agit de l'appliquer.

Lorsque de telles études portent sur des systèmes sociaux (des institutions, des groupes) ou sur un système humain (un sujet), les étapes 2,3 et 4 ne peuvent s'effectuer sur et dans le réel, car des essais «pour voir» ne seraient pas sans effets durables. De plus, comme il s'agit d'essais qui veulent étudier le fonctionnement d'un système dont on ne connaît pas encore les réactions, les appliquer dans le réel serait jouer à l'apprenti sorcier. Ces trois étapes devront donc se réaliser en utilisant des simulations théoriques ou formelles (exploitation d'algorithmes) ou des simulations théâtrales (jeux de rôles) ou imaginaires (scénarios).

5. Négociation de l'implantation : processus par lequel on aboutit à répondre aux questions : quelle modification les acteurs acceptent d'envisager, à quelles conditions, et comment les acteurs peuvent effectivement l'implanter ?

6. Implantation de la modification. Le processus de l'implantation comporte, la plupart du temps, une évaluation et des adaptations de la modification projetée;

7. À partir de la chronique, analyse et évaluation de la stratégie d'implantation, des adaptations effectuées à cette stratégie et de la réalisation effectuée. C'est lors de l'analyse de la chronique que des inférences causales sont effectuées afin d'avoir une vue compréhensive qui permettra la phase suivante : l'évaluation;

8. Évaluation de l'intervention. Cette évaluation suit le principe de l'évaluation par comparaison interne : la confrontation des objectifs avec les réalisations étant donné les conditions accordées;

9. Rédaction et communication du rapport (couvrant l'analyse de la demande, le choix de la modification, l'analyse et l'évaluation faites à partir de la chronique); et enfin,

10. Examen des effets de la communication du rapport sur le système demandeur. En effet, lorsque l'intervention proprement dite est terminée, la communication du rapport peut encore avoir des effets, parfois inattendus et indésirables, sur le fonctionnement du système. Il importe donc, si l'intervention veut être source de connaissances pour des actions futures, de prolonger le contact avec le système demandeur afin d'évaluer l'impact du rapport.

Quel que soit le raffinement dans l'identification des étapes, leur importance variera selon l'accent mis par le chercheur sur son implication et sur l'implication de la population ou de l'organisation demanderesse, ainsi que sur l'enseignement ou les connaissances qu'il(s) souhaite(nt) retirer de l'intervention. À cet égard, on ne saurait trop insister sur l'importance que revêt la chronique de l'intervention et son analyse. Dans la mesure où l'on souhaite que l'intervention ait un impact autre que strictement local et momentané (le problème posé est résolu, sans plus), c'est-à-dire lorsqu'on veut en valider et en diffuser les retombées conceptuelles et stratégiques, la rédaction de la chronique et son analyse (qui constituent les *inscriptions* d'une telle recherche) requerront autant de soins que l'implantation et les étapes qui auront permis de la préparer. Dans une perspective de recherche pour l'éducation, cela est essentiel.

Enfin, le pédagogue intervenant ne doit pas se leurrer sur son statut (acteur-consultant, participant-observateur) et sur ses responsabilités (auteur-traducteur) s'il veut ménager sa crédibilité. En conséquence, ou bien l'intervenant énonce qu'il reste un consultant pourvu de savoir et qui retirera de l'intervention un savoir dont il pourra seul bénéficier (publications, notoriété et promotions), ou bien l'intervenant est au service de l'intervention et il ne peut que participer au gain collectif. Dans ce dernier cas, il conviendrait que les institutions dont ces chercheurs dépendent, reconnaissent au bénéfice du chercheur ainsi impliqué la valeur du gain et du produit collectif.

2.3 L'innovation et sa dissémination

Le mot innovation a subi une forte dévaluation ces vingt dernières années par suite de la croyance un peu naïve, même lorsqu'elle est ministérielle ou syndicale, que la «nouveau» et le «renouvellement» apporteront des solutions à des problèmes parfois séculaires, comme celui de l'échec scolaire. Tout changement dans une école, dans une classe est trop rapidement appelé innovation, même lorsqu'il ne s'agit que d'importer dans un lieu des pratiques vieilles de quatre siècles ou implantées depuis

longtemps dans d'autres écoles. Aussi, de nombreuses «innovations», qui ne sont que des introductions de changement, des réformes ou des modifications de pratiques tireraient avantage, pour être efficaces, à suivre un modèle de recherche comme ceux qui ont été exposés aux points 2.1 et 2.2. L'innovation (ou l'invention, la création) pédagogique est essentiellement une démarche personnelle réalisée par l'innovateur et pour lui dans son propre milieu : l'innovateur n'est pas un réformateur de la pratique des autres.

L'innovation pédagogique est un genre de production qui pourrait aussi être qualifiée de recherche à condition qu'elle s'astreigne, mais après coup, à des exigences semblables à celles qui sont attendues de la recherche-développement (voir 3.4). En son principe, la création pédagogique (invention d'un matériel, d'une stratégie) ne supporte pas la démarche du développement. Cette démarche anéantirait la possibilité de l'intuition créatrice, de l'improvisation pédagogique. Mais si on ne s'y soumet pas une fois la création effectuée, aucune nouvelle connaissance ne peut en émerger et la diffusion sera localement et temporellement limitée.

Comme l'indique l'organigramme de la figure 7.4, une fois l'innovation réalisée, un récit doit en être fait, une chronique doit être constituée : il faut une trace de l'innovation, une inscription pour l'analyser et y réfléchir. Bien que rétrospective, la chronique devra, dans ce cas aussi, être la plus fidèle possible au déroulement des événements, sans y ajouter une continuité qui était absente, puisqu'il n'y a pas eu de création sans rupture entre deux rationalités.

Une fois la chronique disponible, le chercheur, qu'il soit l'innovateur lui-même ou un tiers, peut l'analyser et en construire la compréhension. S'il souhaite que l'innovation soit diffusée, il devra analyser le problème auquel l'innovation est une réponse, y compris le contexte qui a favorisé la rupture, ainsi que le coût des investissements émotionnels et matériels consentis pour réaliser l'innovation. En effet, toute autre personne ne sera intéressée à investir dans un changement que si elle peut reconnaître le problème à la base de l'innovation, estimer la pertinence de cette réponse dans son propre contexte de travail et anticiper les coûts qu'elle se sent capable d'assumer. C'est le plus souvent parce qu'il n'y a pas de chronique satisfaisante pour qu'une recherche (réflexive ou évaluative-adaptative) puisse être entreprise que les innovations restent sans lendemain.

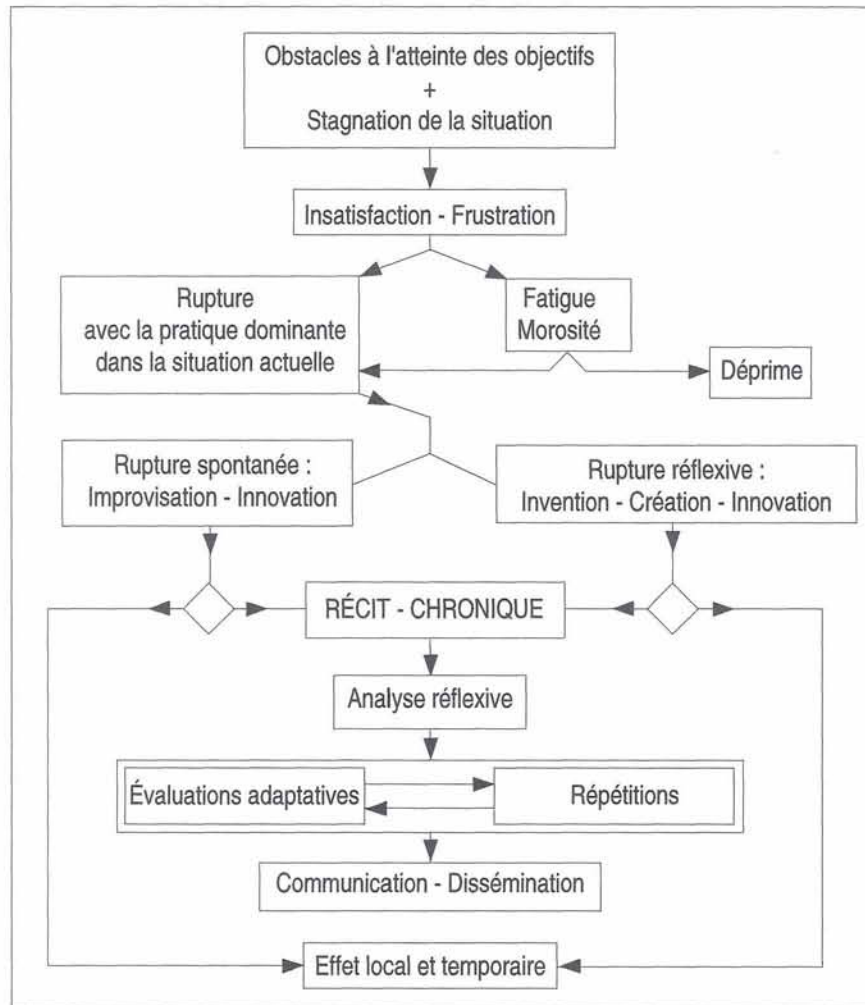


Figure 7.4 : L'innovation comme recherche et sa diffusion.

3 La recherche de développement

La recherche de développement peut prendre trois formes : le développement de concept, le développement d'objet ou d'outil et le développement ou perfectionnement d'habiletés personnelles en tant qu'outils professionnels.

3.1 Le développement de concept

Le développement de concept consiste à chercher à quelles applications ou à quels développements d'outils peuvent conduire certains énoncés théoriques. Il s'agit, en fait, de vendre une idée après l'avoir opérationnalisée, c'est-à-dire après l'avoir transformée en outils, en moyens, en procédures applicables dans des situations concrètes. Dans bien des cas, ce genre de recherche ne peut se faire par le théoricien, car la transformation d'un énoncé théorique en un développement technologique ou pratique implique souvent un abandon de la structure complète de la théorie : c'est une analogie entre la théorie et une situation pratique qui conduit le chercheur à imaginer l'usage qui pourrait être fait de la théorie. Mais de l'imagination de cet usage à sa mise en marché, il y a souvent un écart dans lequel la théorie voit certains de ses constituants modifiés. Une fois l'usage possible imaginé, un double travail reste à réaliser : d'une part, le développement, soit conceptualisation de l'usage de la théorie, modélisation, simulation, essais de prototypes et corrections multiples; d'autre part la recherche d'un marché et la création du besoin, ce qui peut se faire en s'inspirant de la recherche-action aux enjeux politiques (il s'agit bien d'un changement de comportements – acheter le concept – à produire chez autrui).

3.2 Le développement d'objet

La seconde forme de développement vise la solution de problèmes formulés à partir de la pratique quotidienne en utilisant diverses théories élaborées par la recherche nomothétique. Dans ce cas, la recherche appliquée est efficace dans la mesure où elle apporte des solutions aux problèmes et, pour y arriver, les théories auxquelles on recourt ne font que servir la recherche de solutions. Cette démarche intéresse le champ de la didactique, et c'est le domaine, avec la technologie de l'éducation, où l'on rencontre surtout ce type de recherche.

Le développement d'un objet pédagogique, qu'il s'agisse de procédés, de moyens, d'un milieu, etc., doit suivre un cheminement proche de la résolution de problème s'il veut se faire avec un minimum de rationalité. On commence par l'analyse de l'objet qui répondrait ou bien à l'analyse des besoins chez la population-cible, ou bien à l'analyse du concept que l'on souhaite opérationnaliser. Ensuite, on conceptualise l'objet afin de pouvoir en élaborer un modèle, c'est-à-dire une représentation cohérente des éléments qui doivent le composer et des contraintes auxquelles il doit répondre; tant pour les composants que pour les exigences, il faut établir

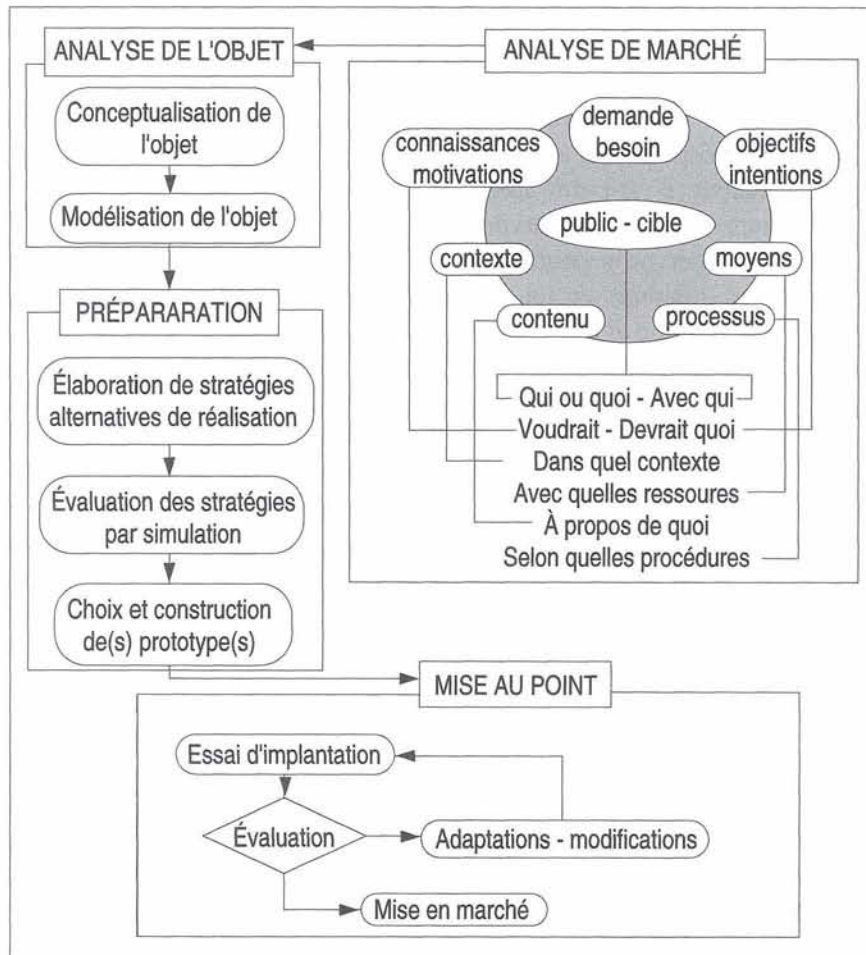


Figure 7.5 : Les quatre étapes du développement d'objet.

un ordre de priorité permettant d'établir la liste ordonnée des éléments que l'on pourrait éventuellement sacrifier advenant des difficultés de réalisation. Une fois cette représentation cohérente établie, on élabore plusieurs stratégies alternatives de réalisation en sachant que le produit final sera sans doute un compromis entre ce qui est souhaité et ce qui est possible étant donné les contraintes. On procède alors à une première évaluation des différentes solutions par simulation mentale et matérielle de modèles réduits ou formels de l'objet, à partir de laquelle se fait le choix du projet retenu. L'anglicisme «design» recouvre les étapes suivies depuis la conceptualisation de l'objet jusqu'au choix du projet. Lorsqu'un projet est retenu, on passe à la construction d'une forme provisoire de l'objet : le

prototype. Enfin, on passe à la mise au point du prototype qui subit une série d'essais (essai, évaluation, adaptation, modification, essai, évaluation, etc.) en situation artificielle (laboratoire, école pilote), puis sur un échantillon de situations réelles. Une fois que les exigences de la mise au point sont complétées, il reste à passer à la phase finale d'implantation, soit la phase de mise en marché.

Une variante de ce type de recherche, dite *collaborative* et privilégiée en didactique, préconise que l'analyse de l'objet et sa préparation s'effectuent en collaboration avec les enseignants, de manière à tenir compte immédiatement de leurs contraintes et de leurs priorités. De plus, la construction du prototype et ses mises à l'essai s'effectuent non plus en situation expérimentale contrôlée, mais directement en contexte scolaire. Le chercheur participe avec des enseignants et leurs élèves à la recherche de procédures applicables en situations scolaires quotidiennes dans des classes ordinaires : il se confronte aux résistances du milieu et compose avec les contraintes qui transforment un modèle idéal en production réaliste.

Chacune des étapes doit être assez poussée avant de passer à la suivante : on n'en doit éluder aucune malgré les tentations. Par exemple, il faut éviter de n'envisager qu'une seule stratégie (celle qui correspond à l'intuition, celle qui semble évidemment la bonne du premier coup) et de ne pas procéder à l'évaluation de plusieurs possibilités. La boucle «essai d'implantation, évaluation, adaptations et modifications» doit souvent être répétée avant que l'évaluation finale n'aboutisse à la décision de proposer le matériel aux utilisateurs. Dans certaines commandites, une forme particulière de cette boucle évaluative est exigée dans le déroulement de la recherche de développement. Elle s'inspire de l'évaluation adaptative-interactive et elle a été schématisée par Stolovitch¹¹ sous l'étiquette L.V.R. pour *Learner Verification and Revision*. Dans certains cas, le processus L.V.R. vise plus que l'adaptation du matériel, il s'agit alors d'en faire la promotion. À cette fin, il faut diffuser le rapport de la mise au point du prototype auprès d'un échantillon de la population-cible.

Le processus L.V.R. (figure 7.5) peut être réalisé selon deux formules. La première, plus classique, consiste à effectuer la vérification auprès d'échantillons d'élèves extraits de la population-cible. Elle produit des résultats statistiques, mais si le prototype n'est pas suffisamment au point

11 Stolovitch H.D., *La vérification et la révision du produit pédagogique auprès de l'étudiant : une technologie intermédiaire*. Montréal, Université de Montréal, 1979. Cité par G. de Landsheere, *La recherche expérimentale en éducation*. Paris : UNESCO, 1982, p. 60-62.

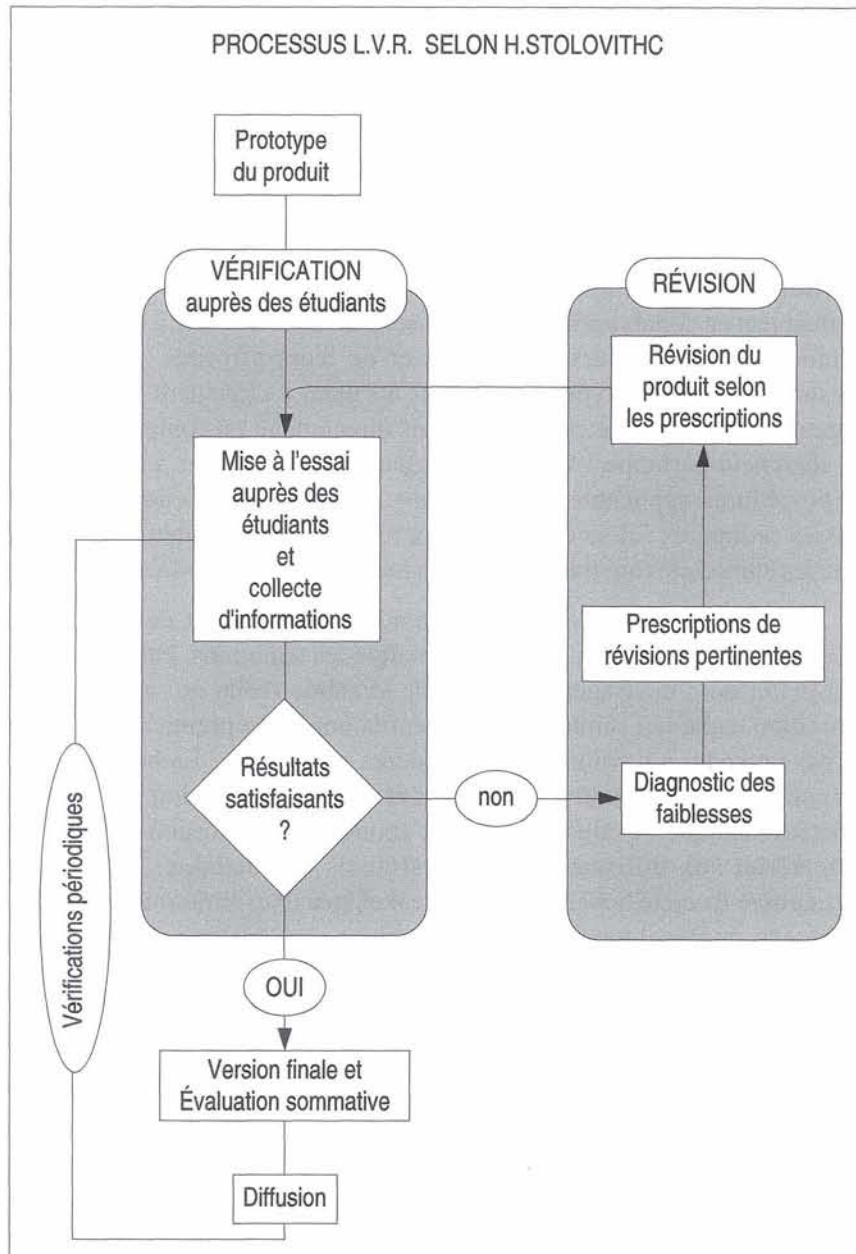


Figure 7.6 : L'organigramme d'une évaluation L.V.R.

avant d'enclencher le processus L.V.R., le chercheur est contraint de reprendre plusieurs fois la boucle de vérification et de révision, ce qui peut «consommer» un nombre considérable de groupes d'élèves qui ne sont pas toujours accessibles.

Afin de pallier cet inconvénient, on peut procéder par *chaînes évaluatives*. Selon cette formule, chaque boucle L.V.R. est appliquée sur un sujet différent, la chaîne comprenant de trois à cinq sujets. Dans une chaîne, on applique au sujet numéro 2 une première révision du prototype, révision obtenue à la suite de la vérification avec le sujet numéro 1, puis on applique au sujet numéro 3 une seconde révision du prototype obtenue grâce au sujet numéro 2, et ainsi de suite jusqu'au cinquième sujet. Une chaîne comporte donc de trois à cinq révisions du prototype. Pour tenir compte de l'effet de diverses variables, dont le contexte journalier, nous avons pris l'habitude d'effectuer au moins trois chaînes de cinq sujets et de comparer la série des révisions effectuées, chaque chaîne débutant avec le prototype initial. La formule des chaînes évaluatives apporte des informations de types plus cliniques, elle permet souvent de mieux tenir compte des perceptions des élèves et d'individualiser, dans la mesure du possible, la présentation du matériel, mais, bien qu'économe en sujets, elle consomme plus de temps que la formule classique. Le choix d'une des deux formules dépendra du temps disponible, du nombre de sujets accessibles, de l'information souhaitée et, si on ne peut l'influencer, des exigences du commanditaire.

3.3 Le perfectionnement des habiletés personnelles comme développement d'outils professionnels

Plus d'une recherche évaluative ou de développement ont fondamentalement une visée ontogénique : il s'agit pour le praticien de développer ses habiletés et ses connaissances en tant qu'instruments d'intervention professionnelle. La démarche suivie s'inspire directement de ces deux types de recherche, c'est-à-dire d'une recherche évaluative pour ce qui correspond à l'analyse de marché et, ensuite, d'une recherche de développement. Dans cette dernière, l'analyse ne porte pas sur un objet externe, mais sur des habiletés ou des connaissances qu'il s'agit de s'approprier selon diverses possibilités. Comme dans l'innovation, et à l'instar de ce qui a été discuté à propos de l'élaboration d'une praxis (au chapitre 2, paragraphe 4.1.), deux exigences doivent être respectées lors de la phase évaluative et lors du développement : la constitution d'une trace primaire de la démarche, le plus souvent accompagnée d'une chronique, et le recours à un tiers-témoin analyste.

Pour que les analyses d'une pratique (c.-à-d. d'une habileté que l'on développe) puissent avoir quelque pertinence, il faut que la trace obtenue soit une représentation la plus semblable possible à l'événement tel qu'il s'est produit dans le temps et dans l'espace, dans sa complexité et avec le jeu des divers acteurs. La trace primaire doit être captée en situation, c'est-à-dire au cours du déroulement de l'action; elle doit provenir de la classe dans laquelle l'enseignant s'exécute ou d'un atelier-studio au sein duquel on se prépare à l'action et on apprend en l'essayant. Techniquement, il est maintenant possible d'obtenir de telles traces sans perturbations majeures de l'action : l'enregistrement vidéoscopique peut se faire à distance (zoom de 14 x), sans éclairages spéciaux (comescope travaillant avec 2 lux) et avec une excellente fidélité, du moins pour l'utilisation qu'on veut en faire. Lorsque l'enregistrement vidéoscopique n'est pas possible (par exemple, dans les cas où l'action déborde la classe), une chronique des événements doit être écrite. Cette chronique est enregistrée et fixée le plus vite possible après le déroulement de l'action, elle doit respecter au mieux les formes de l'action et n'avoir que les apparences d'un scénario : on y raconte comment les choses se déroulent, dans leur simple succession et dans leur décor et sans inférer des relations causales entre les événements. C'est, ensuite, l'analyse réflexive de la chronique qui permet la construction d'une histoire dans laquelle les actions s'enchaînent grâce à leurs motifs; mais la chronique est seulement le rapport de la succession des actions.

En ce qui concerne le choix du tiers-témoin, la première personne à envisager est un autre praticien, et pas nécessairement un senior. Même un collègue plus jeune peut poser de bonnes questions et suggérer des solutions pertinentes à la situation examinée. Le meilleur analyste du travail d'un artiste, lorsqu'il s'agit non pas de critiquer mais d'améliorer, doit être un comparse (mais pas un complice) : pour un acteur, c'est un autre acteur, le metteur en scène ou un autre artisan du spectacle. La meilleure analyse provient de quelqu'un qui fréquente la coulisse, c'est-à-dire de quelqu'un qui est au courant des contraintes qui délimitent l'action. Encore faut-il trouver un collègue qui ait la disponibilité et l'intérêt pour une telle recherche. Aussi ce sera sans doute plus souvent un formateur (le tuteur ou le directeur de recherche) ou un autre chercheur qui se présentera comme analyste. Mais, plus encore que dans l'élaboration d'une praxis, le tiers-analyste devra, dans la recherche ontogénique, être un co-analyste, un provocateur d'analyse plus qu'un énonciateur d'analyse, ce qui n'est pas simple lorsque ce rôle est joué par un formateur.

LES DÉMARCHES EMPIRISTES D'UNE RECHERCHE NOMOTHÉTIQUE

- 1 Quelques axiomes empiristes de la recherche nomothétique
- 2 Les stratégies de la recherche empiriste
- 3 L'induction d'hypothèses : la recherche exploratoire
 - 3.1 *La stratégie statistique*
 - 3.1.1 L'instrumentation et l'analyse des données
 - 3.1.2 La synthèse des données et la production de résultats
 - 3.1.3 L'interprétation des résultats : induire et valider des hypothèses
 - 3.2 *La monographie*
- 4 La déduction d'hypothèses : la recherche vérificative
 - 4.1 *La vérification de la validité logique*
 - 4.2 *La vérification de la validité empiriste*
 - 4.2.1 La confirmation
 - 4.2.2 La vérification infirmative
 - 4.3 *La décision dans une recherche vérificative*
 - 4.3.1 La détermination préalable des seuils
 - 4.3.2 Le contrôle des aléas d'échantillonnage
 - 4.3.3 Le contrôle des variables parasites
 - 4.3.4 Les incertitudes de la décision dans la vérification confirmative
 - 4.4 *L'étude de cas vérificative*

Dans ce chapitre le terme *empiriste* a été particulièrement retenu pour souligner qu'il s'agit d'une perspective rationnelle de développement des connaissances et non pas du développement spontané, expérientiel d'un savoir personnel. Le terme empirique est souvent utilisé dans les écrits méthodologiques à la place du mot empiriste et avec le sens de ce dernier («propre à l'empirisme ou partisan de l'empirisme», dictionnaire *Robert*), alors que son sens premier en français y est quasiment opposé : qui s'appuie sur l'expérience et non pas sur des données scientifiques ou rationnelles.

1 Quelques axiomes empiristes de la recherche nomothétique

La recherche nomothétique n'a de sens que si l'on accepte quelques axiomes, c'est-à-dire quelques définitions de base indémontrables qu'il faut respecter comme telles pour pouvoir poursuivre cette activité.

1.1 La recherche nomothétique a pour objectif de produire des connaissances, c'est-à-dire des énoncés théoriques qui ont la prétention de s'appliquer de la manière la plus générale possible. Ces énoncés, appelés «*lois scientifiques*», établissent des relations entre faits, événements, caractéristiques ou traits d'un sujet. À condition de respecter certaines restrictions, la relation énoncée est censée stable et applicable à l'ensemble du champ pour lequel elle a été élaborée. La loi a donc une prétention générale et des ambitions de prédiction pour un univers déterminé. Les restrictions dont il faut tenir compte sont reliées au caractère fondamentalement hypothétique des énoncés nomologiques (voir ci-dessous 1.3).

1.2 L'utilisation du terme *loi* conduit à une fausse conception, comme si, en analogie avec le droit, la loi, une fois formulée, s'appliquait et contraignait tous les objets du domaine en question, ce qui n'est pas le cas. En effet, ce ne sont pas les connaissances du chirurgien qui font que le coeur des humains se situe habituellement au même endroit du corps; c'est la régularité avec laquelle les anatomistes ont découvert le coeur approximativement au même endroit qui les a conduits à énoncer qu'en principe, ou en règle générale, les chirurgiens pouvaient compter l'y trouver. D'une part, il conviendrait peut-être de privilégier le terme *principe* à la place du terme *loi*; d'autre part, si c'est la raison qui construit les règles à partir des observations, la raison doit rester soumise au contrôle des faits d'observation, car la raison seule qui infère à partir de quelques observations, pourrait en dire des choses que d'autres observations ne confirment pas.

1.3 Toutes les connaissances scientifiquement établies sont hypothétiques : elles n'ont de valeur que dans la mesure où une nouvelle observation ne vient pas les invalider. Il n'y a pas de théories scientifiques capables d'énoncer une «vérité» absolue; les énoncés scientifiques ne sont pas «vrais», ils ne sont que valides, et leur validité est relative à un domaine d'application dont aucun chercheur ne peut être certain d'avoir définitivement cerné les frontières. À moins que l'objet, dont la théorie parle, soit un objet construit et que ses modes de construction soient immuables et insensibles aux variations de l'environnement, une relative incertitude subsistera toujours quant aux limites de la validité des énoncés théoriques. Les énoncés nomologiques sont hypothétiques par leur caractère

- provisoire (les choses peuvent changer tant qu'elles ne sont pas mortes);
- conditionnel (c'est seulement si certaines conditions sont respectées qu'une relation peut être observée);
- relatif (les énoncés ne sont valables que du point de vue à partir duquel les observations ont été faites); et
- conjectural (l'énoncé est un pari probable inféré sur la base d'une partie raisonnable de toutes les observations qu'il aurait fallu pouvoir faire pour obtenir une certitude).

C'est la reconnaissance de ce caractère hypothétique des énoncés qui fait de la recherche une entreprise finalisée (c.-à-d. orientée par des finalités), mais sans fin.

La relativité des énoncés nomologiques semble paradoxale par rapport à leur visée générale, du moins pour leur domaine d'application. D'où la tentation des théoriciens de passer d'une perspective nomologique à une perspective idéologique : ils construisent des énoncés supplémentaires, qui sortent les premiers énoncés du contexte qui les limite, afin de permettre à ces premiers énoncés de prétendre à la vérité absolue alors qu'ils n'ont qu'une valeur relative. Jean-François Kahn montre très bien par de nombreux exemples que : «... la vérité qui découle (de la recherche scientifique) est relative. Or elle ne s'accepte presque jamais comme telle. Sinon il n'y aurait ni aristotélisme, ni platonisme, ni positivisme, ni marxisme, ni fidéisme, ni libéralisme : il est dans la nature d'une vérité relative de se croire absolue¹.»

1.4 La démarche nomothétique est normalement polémique et contestataire. Si l'on accepte à la fois la prétention générale et le caractère

1 Kahn J.F., *Esquisse d'une philosophie du mensonge*. Paris : Flammarion, 1989, p. 288.

hypothétique des énoncés théoriques, la recherche qui se donne pour fonction de faire progresser les connaissances, autrement dit d'en savoir plus et de manière plus précise, ne peut le faire qu'en maintenant le doute sur les énoncés déjà prononcés. Ce doute est d'autant plus justifié que l'on vient de voir que les énoncés, une fois prononcés, ont tendance à se constituer en doctrine plutôt qu'en théorie, à se prendre pour absolus alors qu'ils ne sont que relatifs. Il ne s'agit pas nécessairement de prétendre que ces énoncés sont faux; il s'agit au moins d'envisager qu'ils ne sont peut-être valides que dans des conditions (caractère conditionnel), d'un point de vue (caractère relatif) et pour un domaine (caractère conjectural) plus restreint que ceux pour lesquels ils ont été énoncés ou pour lequel ils prétendent s'appliquer. À la différence de la recherche policière ou judiciaire qui doit «prouver hors de tout doute raisonnable» (la certitude absolue étant impossible), la recherche nomothétique doit *éprouver*. Le policier, le ministre, le professionnel souhaitent garantir au public qu'un chercheur a pu «administrer la preuve» de la valeur d'un jugement ou d'une décision, alors que les chercheurs savent qu'ils ne peuvent que *soumettre à l'épreuve* des faits empiriques les prétentions d'un énoncé hypothétique.

1.5 En éducation, la recherche a toujours comme fonction de mieux décrire, comprendre, expliquer ou formaliser une situation éducative. Or, cette dernière n'est pas un pur concept à l'instar de l'âme, de la pensée ou de la langue. La situation éducative est un objet concret, humain, dont l'existence et le fonctionnement tiennent aux significations et aux valeurs attribuées par les acteurs aux échanges symboliques (les gestes, les mots, les écrits) qui constituent cette situation. En conséquence, les premières étapes d'une recherche pédagogique (c.-à-d. d'une recherche qui, en priorité, instrumente les acteurs et fonde leur action) devront comporter une description du problème tel qu'il se présente sur le terrain et une explicitation de la manière dont il est appréhendé par les acteurs : leurs catégories et leurs préconceptions, leurs attentes et leurs contraintes, leurs urgences et leurs valeurs (y compris les valeurs négatives : leurs refus). Par ailleurs, aucun chercheur s'intéressant à l'éducation ne le fait sans un rapport particulier à ce domaine. Dès lors, sans avoir explicité ses modes d'appréhension, le chercheur ne pourrait pas analyser ses choix théoriques et méthodologiques pour tenir compte des biais qui en découlent, comme il ne pourrait pas les faire accepter par les utilisateurs de sa recherche. L'explicitation des modes d'appréhension du chercheur et des acteurs devra donc faire partie, avec la description du problème *in situ*, des étapes préalables à toute recherche empiriste au bénéfice de l'éducation.

2 Les stratégies de la recherche empiriste

Deux grandes stratégies sont utilisées dans la recherche empiriste : la stratégie statistique (descriptive et inférentielle) d'une part, la stratégie monographique d'autre part. Pour diverses raisons, elles sont souvent opposées alors qu'elles pourraient être complémentaires. Toutes deux ont pour objectif de produire des énoncés hypothétiques et peuvent prétendre aux différents niveaux nomologiques: la description (l'état de la situation), la compréhension (l'origine de la situation, la séquence des événements qui y ont conduit), l'explication (les règles du fonctionnement qui permettent de prévoir ce que la situation deviendra) ou la formalisation (les principes de base à l'oeuvre dans toutes les situations semblables).

Dans la stratégie statistique (au sens large de statistique), on tente de découvrir les traits qui sont communément partagés par le plus grand nombre d'individus, les exceptions formant une classe à part, dont il s'agit alors de décrire les traits spécifiques. Lorsqu'on ne peut examiner tous les individus, la possibilité d'établir des énoncés généralisables est recherchée par la loi des grands nombres : si on peut examiner un grand nombre d'individus avant d'en tirer une inférence, la probabilité d'encore trouver une exception s'en trouve d'autant réduite. La question devient alors celle que tous les novices de la recherche posent : quelle doit être la taille minimale des échantillons, étant entendu qu'il est souvent difficile, sinon impossible, de travailler sur la totalité de la population à propos de laquelle on veut produire les énoncés. La réponse tient du compromis : étant donné la relation qui existe entre la précision d'une inférence (le résultat sera X plus ou moins Y), la probabilité de commettre une erreur dans l'inférence (la chance de se tromper) et la taille de l'échantillon, le chercheur devra choisir ce qu'il sacrifie. Pour une taille donnée d'échantillon, l'inférence sera plus précise (Y tend vers zéro) à condition qu'il accepte une plus grande probabilité d'erreur, à moins que, inversement, il tolère une fourchette de prédiction plus large afin de réduire la probabilité d'erreur. Mais s'il souhaite une grande précision et une faible probabilité d'erreur, la taille de l'échantillon devra se rapprocher de la taille de la population, puisque la seule inférence qui puisse être précise et exacte est la description de la population totale. Lorsqu'on peut s'appuyer sur la loi des grands nombres et sur un échantillonnage aléatoire, la stratégie statistique renvoie donc aux techniques d'échantillonnage, en ce qui concerne la source des données, et aux techniques de l'inférence statistique pour le traitement des données afin d'établir la précision et la probabilité d'erreur des estimations. Dans les cas où l'échantillonnage aléatoire n'est pas possible, le transfert des

conclusions doit être limité aux groupes d'individus qui partagent les mêmes caractéristiques que les individus qui ont fourni les données, ce qui impose un choix rationnel et bien documenté des informateurs. Notons que l'objet de l'inférence peut être de deux ordres : soit le nombre d'individus qui possèdent une caractéristique commune (par exemple, le pourcentage de séropositifs dans une population), soit la valeur moyenne d'un trait manifesté par tous les individus (par exemple, le degré d'instruction ou le niveau moyen de scolarité d'une population). La différence ne relève pas de l'échantillonnage, elle dépend de la qualité des instruments permettant d'obtenir des informations pertinentes pour faire ces inférences.

Dans une stratégie monographique, la généralité des conclusions est recherchée non pas par l'abondance des informations obtenues sur un seul sujet, mais par la réduction de la totalité des informations aux informations essentielles ou fondamentales. Ces informations seraient communes à tous une fois qu'on les a dégagées de ce qui est particulier au sujet qui a fourni l'information et de ce qui relève du contexte dans lequel le sujet se trouvait à ce moment-là. Dans l'étude de cas ou dans l'étude clinique, le chercheur essaye d'obtenir le plus possible d'informations sur le sujet de la recherche, sur sa situation, sur son contexte et son histoire, afin de disposer de toutes les informations nécessaires pour pouvoir y distinguer les données nécessaires (fondamentales, transcendantales) et les données contingentes (contextuelles, épisodiques).

Que la stratégie soit statistique ou monographique, la démarche complète de construction des énoncés comporte d'abord un mouvement inductif ou exploratoire qui vise à générer des hypothèses à partir des faits, puis un mouvement déductif ou vérificatif ayant pour but de mettre les hypothèses à l'épreuve de nouveaux faits. Les deux mouvements sont solidaires, le second ne pouvant précéder le premier et le premier ne pouvant être complet sans le second. Cependant, ce sont souvent des chercheurs différents qui effectuent, les uns, l'induction d'hypothèses, les autres, leur vérification parce que les investissements temporels et financiers ainsi que la spécialisation des techniques et des moyens de recherche exigés par ces deux versants d'une même démarche sont rarement disponibles et maîtrisés par un même chercheur. Dans bien des cas, la spécialisation des chercheurs tient à l'histoire de leur formation et de leurs intérêts, mais elle les conduit à s'intéresser à des questions qui n'en sont pas au même niveau nomologique.

En effet, la recherche vérificative qui utilise une tactique de déduction logique et de prédiction avec des critères de décision précis, ne peut

minimalement se faire que si l'on dispose déjà d'une description instrumentale de l'objet à propos duquel on veut vérifier des hypothèses. Cette exigence implique un domaine où la connaissance est relativement avancée (apparemment plus scientifiquement établie) et favorise les outils quantitatifs pour la constitution des données et le traitement des résultats. Par contre, dans une démarche inductive, le chercheur veut construire des hypothèses sur lesquelles l'élaboration d'instruments s'appuiera. Mais en l'absence d'instruments spécifiquement construits pour son objet, le chercheur doit consentir à se donner une culture scientifique et générale bien plus étendue que le spécialiste de la vérification. Ce dernier peut être efficace en ne maîtrisant que le domaine de son objet de recherche et les techniques vérificatives qui y sont admises, alors que pour réussir une induction d'hypothèses, le chercheur doit repérer des communautés de traits, des similitudes de structures qui ne sont pas évidentes s'il ne dispose que des connaissances propres au domaine de l'objet de recherche, s'il ne peut nourrir des analogies par ses connaissances d'autres domaines. De plus, la collecte de données sur l'objet à explorer ne peut se faire sans instrumentation : ou bien le système inscripteur sera tout usage, ou bien il comportera aussi des instruments empruntés à des domaines voisins. La faiblesse de certaines recherches exploratoires tient non pas au bas niveau nomologique de leur domaine ou au caractère qualitatif des données, mais à l'inculture du chercheur lui-même, souvent trop spécialisé et trop attaché à quelques théories.

3 L'induction d'hypothèses : la recherche exploratoire

La recherche exploratoire a pour but de générer des hypothèses, c'est-à-dire d'examiner un ensemble de données afin de découvrir quelles relations peuvent y être observées, quelles structures peuvent y être construites. Elle cherche à voir quels énoncés pourraient être formulés à propos d'un objet problématique. Pour ce faire, elle pose au départ un *postulat méthodologique*, parfois présenté comme hypothèse méthodologique. On ne sait pas à l'avance s'il donnera quelque chose, mais on doit provisoirement l'admettre comme valable pour au moins l'essayer. Ce n'est que si ce postulat ne permet pas de produire des résultats intéressants qu'il sera remis en question avant d'entreprendre une autre recherche. Pour réduire l'incertitude à propos d'un objet de recherche, le chercheur ne peut se contenter de le contempler; il doit *poser que si* on applique de telle manière tel instrument sur tel matériel (autrement dit, si on utilise tel inscripteur), *alors peut-être* on pourra construire des interprétations nouvelles (sous forme d'énoncés

hypothétiques) sur les données ainsi constituées. Le premier choix de postulat portera sur la stratégie d'induction : statistique ou monographique.

Dans l'un et l'autre cas, l'efficacité d'une recherche inductive exige que le chercheur ait examiné l'ensemble des énoncés et des inscripteurs qui ont déjà été appliqués dans le domaine de l'objet problématique afin que son entreprise puisse aboutir à quelque chose de nouveau; sinon elle n'est pas exploratoire, elle est seulement une mauvaise recherche aboutissant probablement à des énoncés banals, inutiles et redondants avec des énoncés produits ailleurs ou antérieurement. La recherche exploratoire a pour but de combler un vide, une lacune dans les écrits à propos de l'objet, soit en recourant à un nouveau système d'inscripteurs, soit en appliquant des inscripteurs connus à un nouveau matériel; elle doit donc au départ avoir fait le tour d'horizon des écrits sur l'objet problématique afin d'en élaborer, comme indiqué à la figure 8.1, le cadre conceptuel et méthodologique.

3.1 La stratégie statistique

3.1.1 L'instrumentation et l'analyse des données

D'une manière synthétique, on dira que, selon la stratégie statistique, la recherche exploratoire étudie des séries d'événements ou des collections d'individus en visant, par abstraction, à formuler des hypothèses induites. La démarche consiste à repérer des similitudes, soit des structures ou des scénarios (patterns) semblables, ou à identifier les différences qui permettent de discriminer des classes, dans des situations proches mais non identiques. La recherche exploratoire peut aussi comporter une modification d'événements ou d'instruments afin d'observer les conséquences de cette modification sur un environnement donné. Dans ce cas, il s'agit d'une «expérience-pour-voir» et c'est en comparant la forme ou le contenu des objets ayant subi soit une manipulation, soit une observation modifiée, que l'on tente de mieux décrire, comprendre ou expliquer l'évolution des objets.

Autrement dit, le cadre conceptuel débouche sur un vaste choix de techniques de constitution des données (voir la partie 3, chapitres 13, 14 et 15); ce choix est tel qu'aucune recherche exploratoire ne peut se passer de la construction d'un cadre méthodologique relié au cadre conceptuel. La préparation d'une collecte de données suscitées et son codage imposent un cadre conceptuel minimal. Le temps dont les informateurs et le chercheur disposent est assez compté pour que le chercheur construise sa liste de thèmes ou de questions à partir d'un modèle, certes incomplet en phase

exploratoire, mais le moins possible, des éléments clefs de l'objet de recherche. Autrement dit, le chercheur doit avoir identifié les boîtes ou les cases qui composent l'objet et dans lesquelles il cherche à mettre des réponses. Dans le cas des données provoquées, le chercheur doit avoir identifié, au préalable, non seulement les cases à remplir (les rubriques ou questions), mais aussi l'éventail des réponses qu'on peut y mettre (les catégories parmi lesquelles les informateurs devront choisir). De plus, pour que le chercheur puisse organiser l'analyse des réponses, le cadre conceptuel devra comporter des indications quant à la structure des catégories possibles : cela exige donc au moins l'élaboration d'une pré-théorie. Si, comme cela devrait se faire pour assurer la crédibilité des données, le chercheur veut utiliser des données suscitées et provoquées autant que des données invoquées, la démarche implique la construction d'instruments, c'est-à-dire de systèmes d'inscription qui permettent l'enregistrement de traces par l'intermédiaire de procédés déduits d'un cadre conceptuel, sinon d'un cadre théorique.

L'analyse de la valeur des données (ou vérification de leur crédibilité) implique que le chercheur adopte une attitude d'inspecteur de police ou de juge d'instruction devant un témoignage : n'accepter les indications d'une donnée (ou source de données) que si une autre la confirme et ne prendre le recoupement de plusieurs données que comme une approximation de la réalité. Cette nécessité suggère de planifier une constitution des données qui permette de croiser des données ayant un statut réellement différent, étant entendu que le contrôle d'une source par une autre identique n'est pas un contrôle crédible. Certains suggèrent alors, dans la mesure du possible, de confronter des données qualitatives avec des données quantitatives, des données de terrain avec des données de laboratoire, des observations et des expérimentations². On peut exiger mieux : chercher à obtenir des données construites sous des contraintes différentes, soit croiser des données invoquées, des données provoquées et des données suscitées. L'analyse de la valeur des données, sous son triple aspect (fidélité, validité, transférabilité), peut conduire le chercheur à reprendre une partie de la collecte des données de façon à améliorer leur valeur; dans certains cas, il devra même corriger son cadre conceptuel avant de recommencer une autre collecte des données.

2 Brewer J. et A. Hunter, *Multimethods Research*, Newbury Park, CA : Sage Publications, 1989.

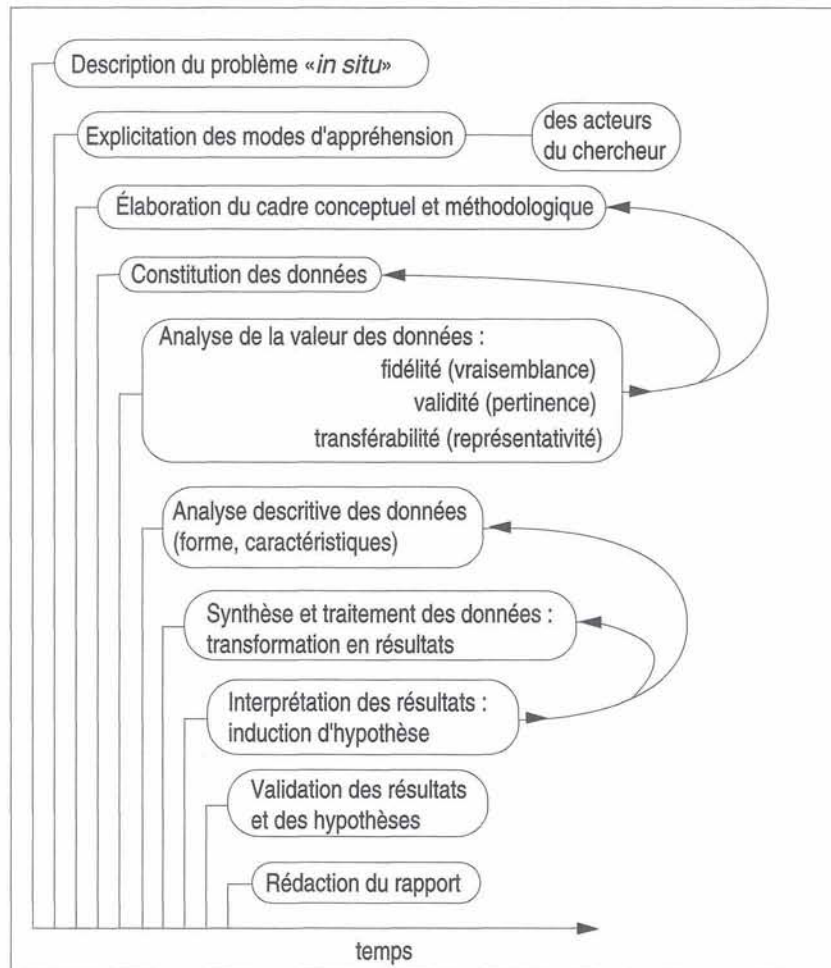


Figure 8.1 : L'induction d'hypothèse : stratégie statistique.

L'analyse descriptive des données se réalise essentiellement par codage puis par comptage, le premier permettant le second. Coder c'est identifier dans le matériel les éléments de réponse aux questions du problème, les marquer, pour ensuite pouvoir plus facilement les repérer, les compter et situer leurs occurrences (qui, quand et comment on en parle). Une fois les comptages effectués, l'analyse se poursuit en examinant les formes que peuvent prendre les données. À cette fin, on utilisera les techniques appelées, à la suite de Tukey³, méthodes d'analyse exploratoire.

3 Tukey J.W., *Exploratory data analysis*. Reading, Mass : Addison-Wesley, 1977.

On trouve une présentation en français de ces techniques dans :

Bertrand R., *Pratique de l'analyse statistique des données*. Sillery, QC : PUQ, 1986.

Ces méthodes exploitent, entre autres, le graphisme : diagrammes en boîtes ou en quartiers de tarte, histogrammes, diagrammes en feuilles, droites et courbes de relations, nuages de points, et autres techniques maintenant incorporées comme outils graphiques à la plupart des logiciels de statistiques pour micro-ordinateurs.

3.1.2 La synthèse des données et la production de résultats

La phase synthétique de la recherche exploratoire peut s'effectuer autant avec des outils qualitatifs, conceptuels, qu'avec des outils quantitatifs. Bien que, dans l'un et l'autre cas, les opérations à effectuer soient semblables, la facilité avec laquelle les techniques numériques produisent des résultats qui « parlent », conduit nombre de chercheurs à privilégier les outils quantitatifs lorsqu'ils peuvent disposer de données quantitatives ou de données qui peuvent être traitées comme telles. La phase synthétique implique généralement une transformation des données à la suite de quoi l'ensemble des données ou des sous-ensembles de données sont remplacées (représentées) par un résultat (par exemple, une moyenne ou un métacode, puis un graphique, une matrice ou un organigramme).

C'est habituellement à partir d'une collection de premiers résultats que d'autres transformations seront effectuées afin de réaliser les opérations de comparaison, de classification, de regroupement, de structuration, de factorisation. Il s'agit donc d'abord de condenser l'information afin de la représenter sous une forme simplifiée; c'est ce que Reuchlin appelle faire des résumés numériques. Ces derniers portent habituellement sur des indicateurs de la tendance centrale (moyenne, médian, mode) et de la dispersion des données (variance, écart type, quantile, quartile). Ces indicateurs, lorsque les données proviennent de la population au complet, ou ces estimateurs, lorsque les données proviennent d'échantillons de la population, permettent facilement (du moins en apparence) de comparer deux ou plusieurs ensembles de données. Ensuite, on peut vouloir examiner si certaines relations n'existent pas entre les données que l'on vient de comparer. Ainsi on pourra examiner, par différentes techniques comme les corrélations, les analyses de la variance et de la régression, si différentes mesures (ou données) obtenues pour une même variable sont interreliées. Par d'autres techniques, comme les analyses de classification en grappes (*clustering*), les analyses factorielles, les analyses de correspondance, on peut chercher quelles relations peuvent apparaître entre différentes observations (des sujets) obtenues pour différentes variables et inversement, quelles relations apparaissent entre différentes variables obtenues pour les différents sujets. Les analyses peuvent ensuite, avec les mêmes techniques,

être reprises à un niveau supérieur d'abstraction afin de voir quelles métarelations pourraient exister entre des relations, c'est-à-dire entre des blocs de données reliées représentés par le facteur exprimant leur relation.

Le travail de synthèse qualitative suivra des chemins semblables lorsque les données sont qualitatives et lorsque le chercheur veut compléter les comptages (fréquences, pourcentages, rangements) par des constructions de sens. Selon différentes formes d'analyse de contenu (voir chapitre 19), par codage, puis par constructions successives de méta-codes plus généraux et de réseaux conceptuels, les chercheurs tenteront de dégager les relations et les formes de relations existant entre les données.

3.1.3 L'interprétation des résultats : induire et valider des hypothèses

Une fois les relations ainsi mises en évidence, l'interprétation théorique reste à faire : les résultats suggèrent des hypothèses qu'il faut formuler. Cette formulation dépend de la capacité du chercheur à comprendre les résultats, à les interpréter. Généralement, à partir de la perception d'analogies, l'interprétation consiste à appliquer aux résultats des concepts extraits de théories antérieures ou extérieures au problème étudié. Cela suppose que le chercheur connaisse ces théories et les résultats à partir desquels elles ont été construites, autrement dit, qu'il sache comment ces concepts avaient été opérationnalisés. Ce ne sont donc pas les résultats comme tels qui produisent les interprétations; ils ne font qu'indiquer l'existence de relations entre variables. L'induction des chaînons théoriques qui formeront une théorie nécessite l'identification des relations indiquées par les résultats; c'est la connaissance des résultats, c'est-à-dire des relations fondant les théories existantes par ailleurs, qui permettra d'expliquer les nouvelles relations observées.

L'induction d'une nouvelle hypothèse à partir de résultats implique un double contrôle : il importe d'abord de valider son énonciation par l'examen du programme de recherche qui y a conduit, avant de vouloir la mettre à l'épreuve dans une recherche vérificative. L'examen du programme de recherche, que Miles et Huberman⁴ appellent la vérification interne (effectuée par les chercheurs eux-mêmes) et externe (confiée à des évaluateurs indépendants) des comptes, doit scruter la possibilité d'interférences venant de biais, d'erreurs ou d'artifices non admis. Il doit aussi examiner

4 Miles M. et M. Huberman, *Qualitative Data Analysis*, Newbury Park, CA : Sage Publications, 1984. La version française est éditée chez De Boeck-Wesmael à Bruxelles, sous le titre : Huberman M. et M. Miles, *Analyse de données qualitatives, Recueil de nouvelles méthodes*, 1991.

dans quelle mesure les traitements appliqués pour mettre en évidence les relations n'ont pas modifié, altéré ou supprimé quelques traits essentiels de l'ensemble des données initiales. En effet, si les résumés numériques sont censés représenter la masse des données, il est probable que cette représentation se fasse au détriment de données marginales dont l'importance est parfois sous-estimée. Comme de plus, ces résumés numériques ne sont souvent que des estimateurs obtenus sur des échantillons, et que l'on sait qu'ils ne correspondent avec les indicateurs de la population que dans les limites d'un intervalle de confiance (la fourchette pour laquelle on a un certain pourcentage de chance de ne pas se tromper), il est essentiel de ne pas publier les résultats avant d'avoir procédé à la vérification des comptes et d'avoir vérifié que les hypothèses induites se tiennent aussi bien sur l'ensemble du matériel recueilli que sur les résumés numériques. En effet, au fil des transformations, les résultats pourraient ne représenter que des exceptions ou ne plus être que des artifices par rapport à l'ensemble des données : c'est donc à vérifier.

3.2 La monographie

Selon les disciplines dont les chercheurs s'inspirent, on appellera monographie (sociologie, ethnologie), étude de cas (éducation, médecine, économie, politique et administration) ou recherche clinique (médecine, psychologie), les recherches qui tentent de mettre en évidence des traits généraux, sinon universels, à partir de l'étude détaillée et fouillée d'un seul cas, qu'il s'agisse d'un individu ou d'une situation⁵. Partant du postulat que tous les êtres humains possèdent une même nature, il serait possible de découvrir ce qui en est le propre en deçà du vernis particulier dû aux différences individuelles et culturelles. Comme l'éducation des enfants est un phénomène universel, chaque cas d'éducation d'un enfant ou d'un groupe d'enfants serait susceptible de fournir des informations précieuses sur ce qu'est l'éducation en général, et sur le rapport de l'éducation aux individus et aux cultures particulières. Pour que de telles études puissent arriver à remplir leurs promesses, deux conditions sont essentielles. D'une part, elles doivent être multidisciplinaires dans la construction du cadre conceptuel, et multimodales dans la constitution des données. Comme il s'agit de

5 Yin R.K., *Case Study Research, Design and Methods*. Newbury Park, CA : Sage Publications, Applied Social Research Methods Series, volume 5. 1984, 1989.

Dufour S., Fortin D. et J. Hamel, *L'enquête de terrain en sciences sociales. L'approche monographique et les méthodes qualitatives*. Montréal : Saint Martin, 1991. Comporte une bibliographie annotée.

recueillir le plus d'informations possibles afin de mieux comprendre le cas que ne le ferait son analyse à partir d'une théorie de référence, il est essentiel d'envisager son analyse à partir de données fournies par des instruments provenant d'un large éventail de théories et de disciplines. D'autre part, une fois qu'ils disposent d'une telle information, les chercheurs doivent accepter de procéder à une réduction du matériel recueilli qui puisse distinguer les données générales et les données particulières. Pour ce faire, une étape préalable consiste à identifier les conceptions du chercheur et des acteurs impliqués dans la situation étudiée : analogiquement, on dira qu'il faut reconnaître autant l'ethnocentrisme du chercheur que l'ethnocentrisme des acteurs, pour pouvoir dégager du matériel les données qui transcendent les cultures, puisque ces recherches reposent sur un tel postulat.

Comme dans toutes les recherches qui utilisent des récits et des témoignages autant que des observations, l'analyse de la validité des données, dans les études de cas, doit aussi porter sur la valeur du témoignage : s'il est légitime que le chercheur analyse ses procédures d'observation, il est aussi légitime qu'il analyse les procédures par lesquelles les témoins lui communiquent de l'information. Comme dans la stratégie statistique, l'analyse de la valeur des données recueillies peut imposer au chercheur de retourner sur le terrain et même de modifier son cadre conceptuel et méthodologique.

L'analyse des données nombreuses et variées obtenues sur un cas conduira à une nouvelle hypothèse à condition que l'ensemble des théories et des disciplines contributives soient interpellées dans cette compréhension. Les facteurs explicatifs issus de chacune ne seront pas envisagés en compétition les uns avec les autres : on ne tentera pas de décider lequel d'entre eux est le facteur le plus important, on tentera de tisser la trame qui relie l'ensemble des facteurs qui peuvent avoir joué un rôle, même si c'est à des niveaux différents. Il ne s'agit cependant pas de fondre les théories dans un amalgame confus et de fusionner les cultures dans un *melting-pot* universaliste indifférencié. La construction d'un nouveau modèle hypothétique du cas implique de reconnaître à la fois le contexte culturel dans lequel il s'est développé et celui à partir duquel il a été analysé (chacun ayant une valeur propre), comme elle exige que le choix des facteurs retenus dans la construction du modèle reconnaisse les différents points de vue et les différents niveaux théoriques à partir desquels une analyse pertinente est possible. La relativité des facteurs aux points de vues théoriques qui les ont élaborés, ne signifie pas que le relativisme (selon lequel tous seraient

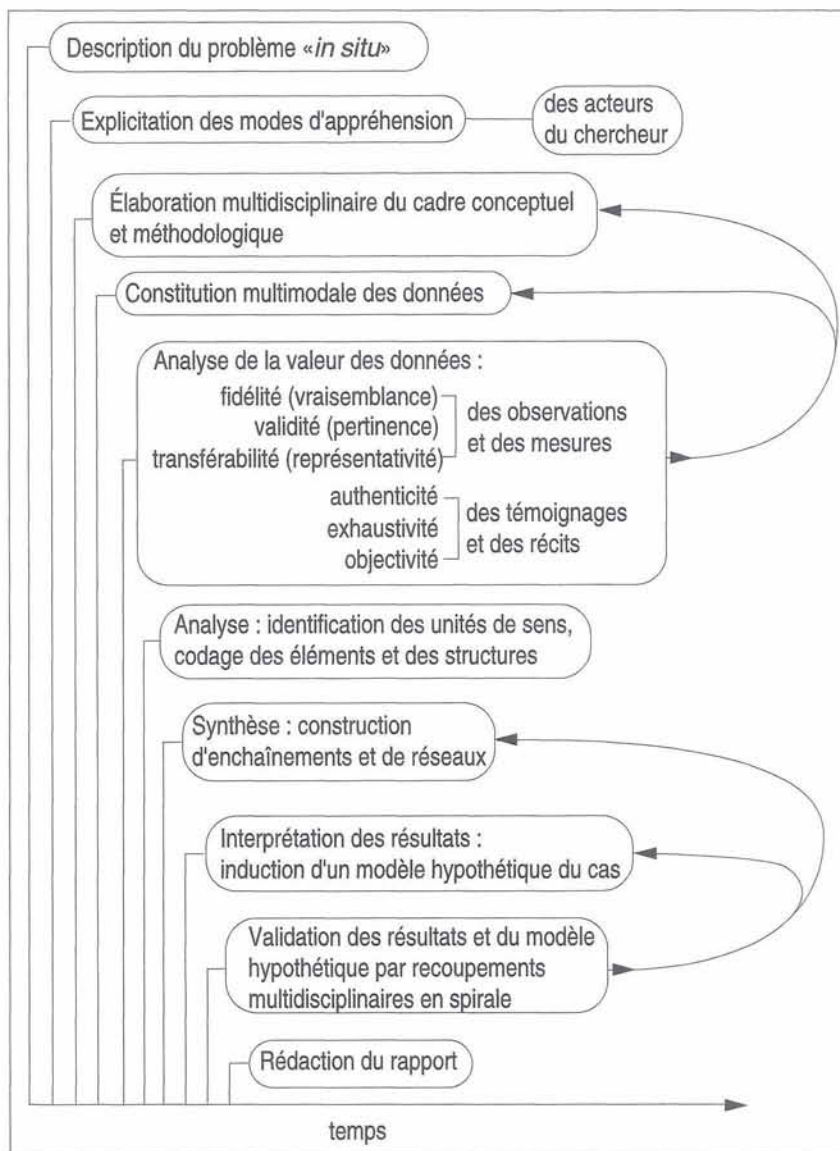


Figure 8.2 : La monographie ou l'étude exploratoire de cas.

également valides) s'applique dans l'évaluation de leur pertinence et de leur contribution à la construction du modèle; il faut pouvoir admettre que dans certains cas, certains facteurs jouent plus que d'autres, certaines explications ou certains ensembles de facteurs rendent mieux compte que d'autres de la situation analysée.

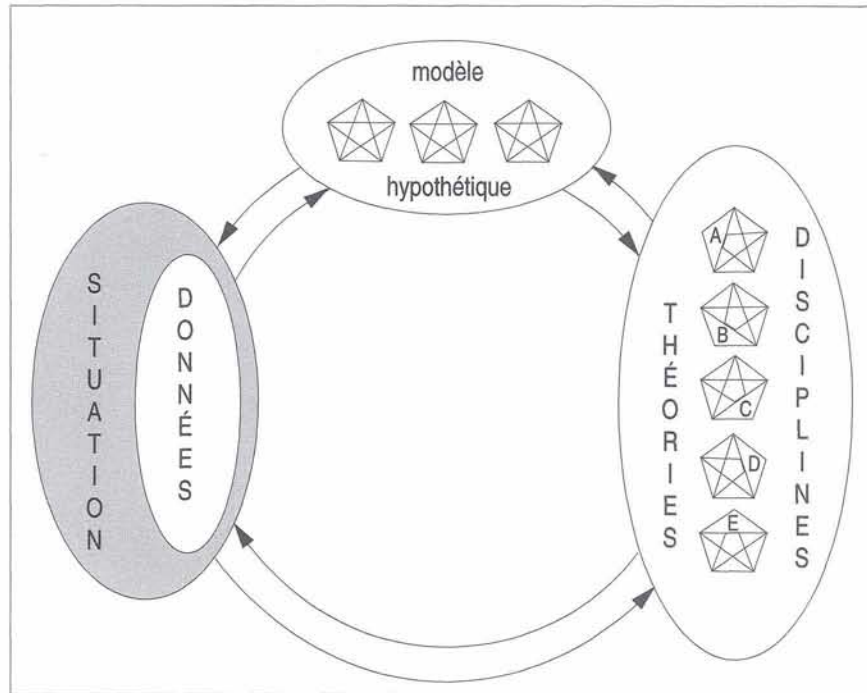


Figure 8.3 : La validation par recoupements multidisciplinaires en spirale.

La vérification de la valeur du modèle induit à partir de l'ensemble des données analysées devra se faire de manière multidisciplinaire : la cohérence du modèle devra être évaluée et corrigée par des aller-retour entre les données et les différents champs théoriques utilisés dans leur constitution et utilisables dans leur analyse.

4 La déduction d'hypothèses : la recherche vérificative

Une fois la vérification des comptes effectuée, la seconde phase de l'élaboration des connaissances peut être envisagée : la phase déductive ou vérificative. Comme indiqué plus haut, il arrive souvent que cette phase soit effectuée par un autre chercheur que celui qui a mené la phase inductive; cette spécialisation ne doit pas faire oublier à celui qui vérifie, que la théorie n'a de sens que par rapport à une situation problématique. C'est la raison pour laquelle, comme indiqué à la figure 8.4, cette phase vérificative commence par un retour à la description du problème initial, qui a suscité la recherche et auquel le cadre théorique tente de donner une réponse, et par l'explicitation des modes d'appréhension propres aux acteurs qui vivent

la situation et au chercheur qui tente de la décrire, de la comprendre ou de l'expliquer. La vérification des énoncés théoriques implique une démarche beaucoup plus rigoureuse que la démarche suivie pour leur élaboration. Cette vérification doit s'effectuer à deux niveaux : celui de la validité logique d'abord, celui de la validité empiriste ensuite.

4.1 La vérification de la validité logique

La première vérification d'une théorie est celle de sa validité logique. En effet, il serait vain de dépenser du temps et de mettre en place des moyens sophistiqués et coûteux pour vérifier empiriquement une théorie qui ne résisterait pas à l'analyse logique. Pour rappel (voir au chapitre 5, paragraphe 3.2.), l'examen de la validité logique comporte quatre exigences que l'on peut présenter sous la forme des questions suivantes :

- a) La suite des énoncés qui forment la théorie respecte-t-elle les règles de la logique (analyse du raisonnement) ?
- b) Les concepts et les relations entre concepts que l'on peut déduire de la théorie sont-ils non contradictoires entre eux, cela tout en respectant la règle de la clôture sémantique (interdiction d'introduire après coup des postulats *ad hoc* pour sauver les apparences) ?
- c) La théorie recoupe-t-elle les explications antérieures et plus simples (apporte-t-elle quelque chose ?) et peut-elle s'exprimer dans un langage commun ou reste-t-elle ésotérique ?
- d) Les rapports entre la théorie et les modèles qui l'opérationnalisent, sont-ils isomorphes, tout comme les rapports entre les modèles et les objets sur lesquels ils portent, sont-ils homomorphes ? Autrement dit, la théorie et ses modèles parlent-ils d'une réalité en la respectant ou parlent-ils d'une fiction ?

Comme les premières boucles de la figure 8.4 l'indiquent, un échec au test de la validité logique impose au chercheur d'interrompre la vérification et d'entreprendre une recherche spéculative ou une autre recherche inductive afin de corriger la théorie. En effet, il n'est guère utile de s'engager dans une vérification empiriste lorsque la théorie n'a pas de validité logique.

4.2 La vérification de la validité empiriste

La validation empiriste d'une théorie a pour objectif de déterminer, non pas sa vérité, mais son domaine ou champ d'application (dit parfois «domaine de vérité»). La question est de savoir à quelle réalité la théorie

s'applique, selon quelles conditions et avec quelles limites. Deux stratégies (courants épistémologiques) se partagent les recherches vérificatives : la confirmation ou corroboration et l'infirmité ou réfutation.

4.2.1 La confirmation

Le modèle confirmatif consiste à (dé)montrer que si l'on applique une des relations (la thèse) extraite de la théorie à un domaine d'application, l'hypothèse ainsi déduite sous forme de prédiction conditionnelle observable se vérifiera. En conséquence, si une épreuve empiriste montre que l'hypothèse se vérifie, alors la théorie dont elle est déduite tient, du moins dans le domaine où elle a été appliquée. Ce modèle se réalise, le plus souvent encore, par expérimentation : il s'agit non plus d'une expérience pour voir, mais pour éprouver. Dès lors elle doit être sans ambiguïté.

Classiquement la relation (entre les variables) énoncée par la thèse est opérationnalisée de telle sorte qu'on puisse observer l'effet d'une variable (ou ensemble de variables) *indépendante* sur l'autre (ou les autres) variable, dite *dépendante*, en comparant (au moins) deux situations (ou groupes) dans lesquelles l'état de la variable dépendante est observé alors que des variations de la variable indépendante sont systématiquement provoquées. Pour que la confirmation soit valide, il importe que les chercheurs prédisent, avant l'expérience, quelles observations de la variable dépendante ils devront faire lorsqu'une variation de la variable indépendante est produite. Dans ce cas seulement, on considérera que l'existence de la relation entre les deux variables est démontrée. Encore faut-il pouvoir observer, sinon mesurer, les variations possibles de chacune des variables.

C'est la nécessité de mesurer de manière non ambiguë les différents états des deux variables en question et de contrôler les autres variables qui pourraient interférer avec leur relation, qui impose de disposer d'une théorie assez complète (ce qui est encore rarement le cas en éducation). En effet, il faut une théorie de l'objet pour pouvoir en construire des mesures quantitatives, et l'on se rend compte que la théorie doit non seulement comporter des énoncés sur les relations entre les deux variables en question, mais qu'elle doit aussi comporter des énoncés sur leurs relations avec, au moins, les autres variables qui interfèrent de manière notoire avec elles. De ce point de vue, bien des recherches qui se présentent comme confirmatives ne le sont pas car, d'une part, leur hypothèse n'est pas déduite d'une théorie, elle n'est qu'une intuition qui serait à enrichir par de nouvelles explorations et, d'autre part, elles ne sont pas capables d'une prédiction précise de l'ampleur (grandeur), du sens (positif - négatif) et de la fréquence des effets (des différences entre groupes) à observer.

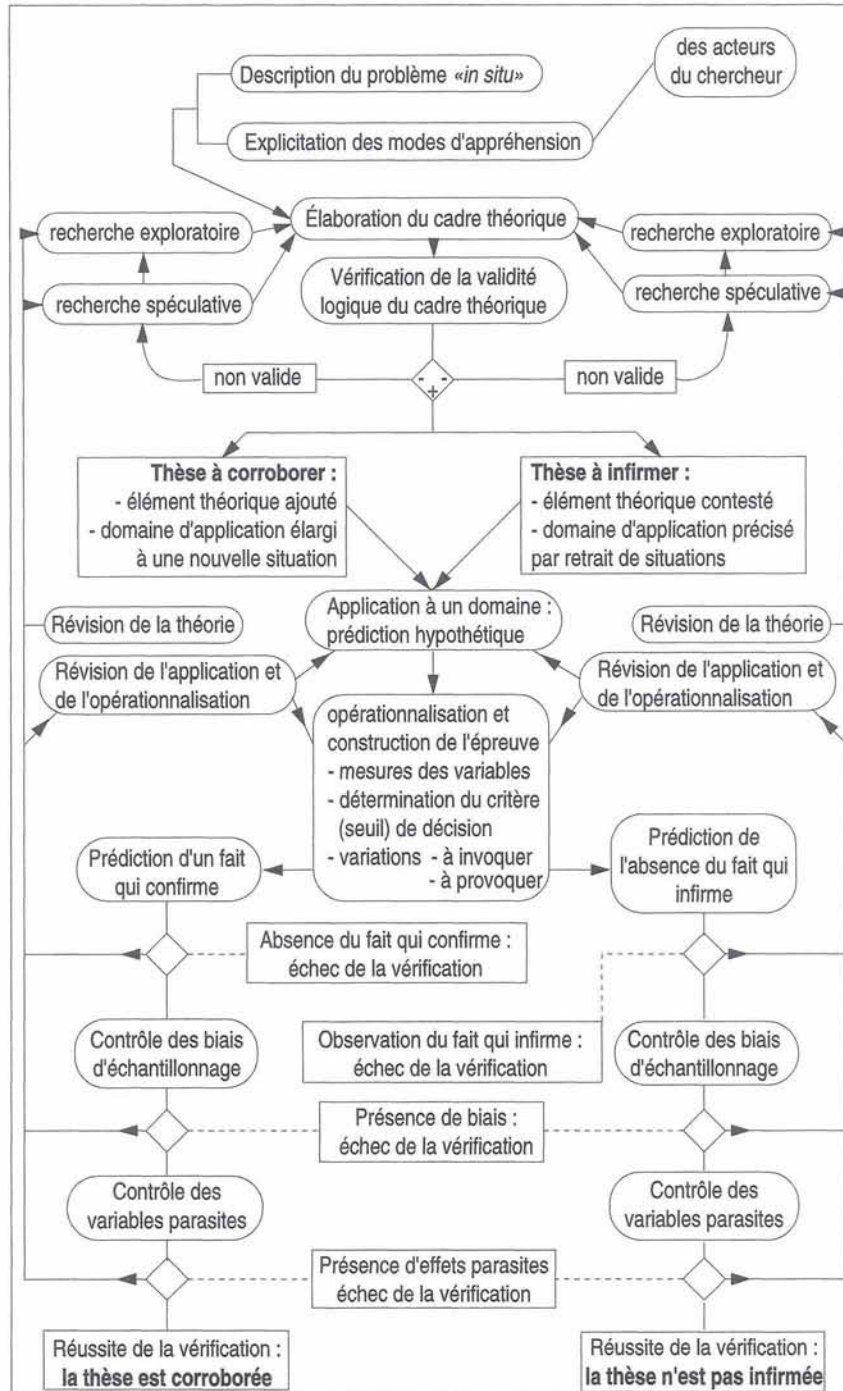


Figure 8.4 : La vérification par confirmation et par infirmation.

Par insuffisance théorique, la comparaison des deux groupes (ou situations) est le plus souvent utilisée sous le mode tout ou rien : dans un groupe, dit témoin ou contrôle, l'état de la variable indépendante est maintenu stable, habituellement à un niveau zéro ou faible, alors que dans le groupe expérimental l'état de la variable indépendante est brusquement modifié. Autrement dit, le traitement est introduit dans le groupe expérimental alors qu'on n'applique rien dans le groupe témoin. On rencontre parfois des expériences où l'état de la variable indépendante est modifié de manière systématique dans le groupe expérimental, c'est-à-dire que le traitement est progressivement introduit selon un programme d'intervention. Dans le cas d'une variation brusque ($0 \Rightarrow 1$), on examine l'effet à un seul moment prédéterminé (le post-test), alors que dans le cas d'une variation progressive, on examine les variations de l'effet à plusieurs moments (une série temporelle) afin de vérifier si les différents états du traitement produisent les différences d'effets prédits au moment prévu.

4.2.2 La vérification infirmative

Le modèle infirmatif consiste à isoler de la théorie un de ses énoncés (la thèse) pour essayer de (dé)montrer qu'il ne s'applique pas avec la généralité qu'on lui accorde, qu'il aurait des exceptions. La recherche consiste alors à appliquer cet énoncé dans un domaine qu'il prétend régir pour observer s'il y a au moins une situation (un groupe) dans laquelle il ne s'applique pas. Si à la suite de la recherche, on peut décider que la situation qui constitue une infirmation (une exception) a bien été observée, il faut réviser au moins l'énoncé testé ou les conditions de son test, sinon revoir l'ensemble de la théorie.

Paradoxalement avec la conception de certains chercheurs, le modèle infirmatif cherche des *résultats non significatifs*. La technique consiste à mettre en place une expérience pour laquelle on peut prédire quelle observation possible infirmerait l'hypothèse déduite de la thèse. Pour tester la validité de la thèse, on va tenter d'observer l'infirmation possible; si elle se produit, la thèse ne tient pas, si elle ne se produit pas, la thèse tient jusqu'au moment où une infirmation sera observée. Une seule infirmation suffit, mais elle doit pouvoir être observable, sinon l'épreuve est une supercherie. La mise à l'épreuve peut se faire de deux façons. La tactique la plus forte est l'expérience critique, soit la mise en évidence d'un fait incontestable qui contrevient à la généralité de la prédiction; mais c'est la moins répandue. L'autre procédure suit la démarche de comparaison entre deux situations, comme dans le modèle confirmatif, mais elle diffère en ce

que, dans ce cas-ci, c'est la situation du groupe témoin qui est visée. On ne cherche pas à montrer que l'hypothèse (l'effet) se produit dans le groupe expérimental où la variable indépendante est manipulée, on cherche à montrer qu'on observe le même effet, ou un effet qui n'est pas suffisamment différent, dans le groupe témoin. Or, ce qui est prédit par l'hypothèse ne peut pas se produire dans le groupe témoin parce que la variable indépendante y est tenue constante ou nulle. Si l'on n'observe aucune différence significative entre les deux situations (celle qui est censée ne pas varier et celle qui, selon l'hypothèse, devrait varier), alors l'hypothèse est infirmée et l'on doit réviser la thèse.

4.3 La décision dans une recherche vérificative

4.3.1 La détermination préalable des seuils

Que ce soit dans un modèle confirmatif ou infirmatif, la recherche vérificative nécessite que des critères précis de décision soient déterminés avant que l'expérience ne commence. En effet, si les critères sont choisis en cours d'expérience ou après le recueil des données, les chercheurs risquent d'être accusés d'établir leurs décisions *a posteriori*, en fonction des données observées et non pas en fonction d'une hypothèse à tester. Ces critères seront le plus souvent exprimés sous forme de seuil ou d'écart minimal pour identifier une variation. La question, assez cruciale, est donc de s'entendre sur ce qui sera considéré comme un indice suffisant de variation tant pour la variable dépendante que pour la variable indépendante. La réponse implique une mesure de variation ou de changement minimal, ce qui ne peut se faire qu'à partir d'une théorie ou d'un consensus entre chercheurs sur un protocole d'expérience. Mais l'histoire des sciences montre que les consensus sont fragiles et que les démonstrations sont plus fortes si les seuils ont été déterminés à partir d'une théorie admise et de postulats reconnus.

Une fois les critères de décision précisés et l'épreuve réalisée, l'étape cruciale de la vérification consiste à décider si les faits observés correspondent ou non aux prédictions. Dès cette étape, on pourra conclure à l'échec de la vérification et reprendre une étude spéculative ou une nouvelle phase inductive si on peut décider que le fait qui confirme ne s'est pas manifesté ou que le fait qui infirme a été observé. Dans les cas où les faits conduisent à une réussite de la vérification, il faut contrôler la possibilité de biais dus à l'échantillonnage et aux variables parasites avant de conclure à la validité de l'épreuve vérificative.

4.3.2 Le contrôle des aléas d'échantillonnage

La décision d'une confirmation ou d'une infirmation ne peut pas seulement reposer sur les statistiques, selon ce qui s'observe trop souvent. Le test statistique de l'hypothèse nulle ne dit rien à propos de l'hypothèse de recherche; il est seulement concerné par l'effet du hasard lorsque des échantillons ont été construits au hasard. Si l'échantillonnage ne s'est pas fait au hasard, le test statistique n'a aucune pertinence. Quelle portée scientifique peut avoir un jeu de langage consistant à dire «si nous avons fait un échantillonnage aléatoire, alors... » lorsqu'il n'y en a pas eu ?

Mais pour quelle raison et dans quel cas utiliser un test statistique ? Préalablement à cette question, pourquoi constituer des échantillons au hasard, ce qui se dit : «*aléatoriser*» le plan de recherche ?

L'aléatorisation (randomization, en anglais) est définie par un triple recours au hasard : lors l'extraction des sujets de la population, lors de leur assignation dans les échantillons et lors de l'affectation des échantillons aux conditions expérimentales. C'est la tactique utilisée par les chercheurs qui se trouvent confrontés à la nécessité de comparer au moins deux groupes de sujets pour évaluer l'effet hypothétique d'une intervention ou d'un traitement auquel un des groupes a été soumis. La manière la plus sûre de vérifier une hypothèse sur un effet consiste à comparer au moins deux groupes (voir chapitre 9) considérés comme identiques avant l'intervention, et à montrer que la différence observée entre les groupes après l'intervention est due à une modification des comportements des seuls individus qui ont subi cette intervention. Une telle démarche comporte un problème quasi insoluble : construire des groupes identiques, du moins pour l'ensemble des traits qui seraient sensibles ou qui pourraient interférer avec l'intervention. Non seulement on ignore souvent quel est cet ensemble de traits, mais en outre, si on le savait, il serait impossible de tester un nombre suffisant d'individus pour constituer des groupes identiques pour l'ensemble des traits considérés.

Cette difficulté peut être contournée si l'on accepte un postulat : la manifestation des traits envisagés se distribue de manière aléatoire dans la population. Si l'on accepte ce postulat, les groupes constitués grâce à une répartition au hasard (c.-à-d. sans qu'un principe n'intervienne) des individus tirés au hasard de cette population, pourront être considérés comme semblables. En fait, par l'*aléatorisation*, c'est-à-dire, en pratique, par une extraction au hasard de la population et par une allocation des sujets au hasard dans les groupes, on construit des groupes qui ne sont pas identiques,

mais qui sont probablement équivalents parce qu'ils sont de bons portraits d'une même population. On dispose alors d'une hypothèse nulle étant donné le hasard ($H_0 \mid A$) qui implique (\Rightarrow) que la différence sera équivalente à zéro ($\Delta \approx 0$) étant donné le hasard ($\mid A$), soit : $H_0 \mid A \Rightarrow \Delta \approx 0 \mid A$. Autrement dit, la différence entre la population et les échantillons est si petite et si peu probable, étant donné le hasard, que les échantillons sont de bons portraits, ou des modèles réduits, de la population. Leur similitude avec la population et entre eux est approximative, la probabilité d'une différence étant simplement due aux aléas de l'échantillonnage.

Une analogie permet de saisir ce qui est en jeu : le sondage. Les sondeurs estiment que les échantillons aléatoires sont de bons modèles de la population. Aussi, ils infèrent le pourcentage de votes de la population pour un candidat à partir du pourcentage observé dans les échantillons. Mais l'inférence est entachée d'une marge d'erreur : il y a probablement une certaine différence entre le modèle (échantillon) et la population. Dès lors on dira que la prédiction est de X (par exemple, 45 %) avec une marge d'erreur (la «fourchette» ou l'intervalle de confiance) de plus ou moins Y (par exemple, 3 et une probabilité d'erreur de Z (par exemple, 5 % ou 1 chance sur 20), Z étant la probabilité que la réalité obtenue à partir de la population soit en dehors de la fourchette prédite à partir de l'échantillon. On comprend que le sondeur vise à optimaliser la marge d'erreur et la probabilité d'erreur : il peut réduire la probabilité d'erreur en élargissant la marge d'erreur. Mais son client (par exemple, un parti politique) préférera sans doute faire un pari un peu risqué sur une prédiction plus précise. En effet, si la marge d'erreur est trop large (par exemple, 45 % des votes plus ou moins 15 % : de 30 à 60 %) la prédiction n'a plus d'intérêt, car l'incertitude de la victoire est trop grande.

Comme les sondeurs, les chercheurs doivent estimer, dès le départ, deux éléments. Ils ont dû décider d'un seuil sur la marge d'erreur : quelle différence entre les échantillons est suffisante pour indiquer une différence avec la population, c'est-à-dire à partir de quelle ampleur la marge d'erreur n'est plus une erreur d'échantillonnage mais une différence. Ensuite, ils doivent décider quel seuil de probabilité associée à l'erreur est acceptable. Par exemple, dire que la probabilité admise est de 5 % ($p = 0,05$) pour un écart «x», veut dire que, étant entendu que l'écart «x» est admis comme ne donnant pas une différence notable, si l'on tire au hasard 100 échantillons de taille «n» de la population, on en trouvera 95 qui lui ressemblent (écart inférieur ou égal à «x») alors que les 5 autres échantillons extraits du même ensemble manifestent un écart plus grand que «x». Si la probabilité associée

à l'écart observé est supérieure au seuil décidé, par exemple 20 %, cela veut dire que dans 20 tirages aléatoires sur 100, on risque de trouver des résultats manifestant un écart plus grand avec la population que l'écart observé. Dans le cas d'une aléatorisation, mais dans ce cas seulement, le test statistique vérifie l'effet de l'échantillonnage sur les résultats étant donné le hasard : il indique la probabilité selon laquelle les échantillons sont de bons portraits de la population. Si la probabilité associée à une différence observée reste inférieure au seuil décidé (par exemple, 5 %), on considère que la différence est due uniquement aux aléas de l'échantillonnage, alors que si elle est supérieure au seuil, on admet que sa probabilité aléatoire est trop élevée pour être due à l'échantillonnage : d'autres variables doivent avoir joué pour créer l'écart observé.

En bref, le test statistique de l'hypothèse nulle n'a de sens que s'il y a eu échantillonnage aléatoire, et il ne fait qu'indiquer si la différence observée est ou non un artefact dû au mode d'échantillonnage. La figure 8.5 illustre cette différence de sens par le schéma d'un plan de recherche comparant deux groupes aléatorisés. Le test statistique porte sur l'hypothèse d'une différence nulle *étant donné le hasard*, « $\Delta \approx 0 \mid A$ » (niveau 11), et non pas sur l'hypothèse d'une différence nulle *étant donné la théorie* : « $\Delta = 0 \mid Th$ » (niveau 8). Or, la décision dans une vérification vise « $\Delta = 0 \mid Th$ », qui est l'*hypothèse alternative* (H_a : niveau 8) à l'*hypothèse de travail* (H_1 : niveau 6), et non pas « $\Delta \approx 0 \mid A$ » (niveau 11) qui résulte du *postulat* (niveau 3) lié à l'aléatorisation (niveau 2).

La probabilité d'une différence due à l'échantillonnage aléatoire lors de la composition des groupes avant l'intervention existe toujours lorsqu'on examine les groupes après l'intervention. L'aléatorisation de la constitution des groupes exige que l'on effectue un test statistique une fois les résultats obtenus : il s'agit de s'assurer que la différence observée au post-test est due aux différences de conditions (niveau 5) auxquelles les groupes sont soumis (c'est l'hypothèse déduite de la théorie sur la variable indépendante : $H_1 \mid Th$, niveau 6) et qu'elle n'est pas due au hasard (niveau 2) qui a été invoqué lors de la constitution des échantillons ($H_0 \mid A$, niveau 3)). Comme indiqué dans la figure 8.5, la décision devra être double (niveau 7). D'abord décider si la différence observée étant donné le traitement est bien suffisante pour affirmer que le traitement a eu un effet ($H_1 \mid Th \Rightarrow \Delta > 0 \mid Th$, niveau 6) et, s'il y a un effet, vérifier s'il n'est pas dû à l'aléatorisation ($H_0 \mid A \Rightarrow \Delta (03 - 04) \approx 0 \mid A$, niveau 11).

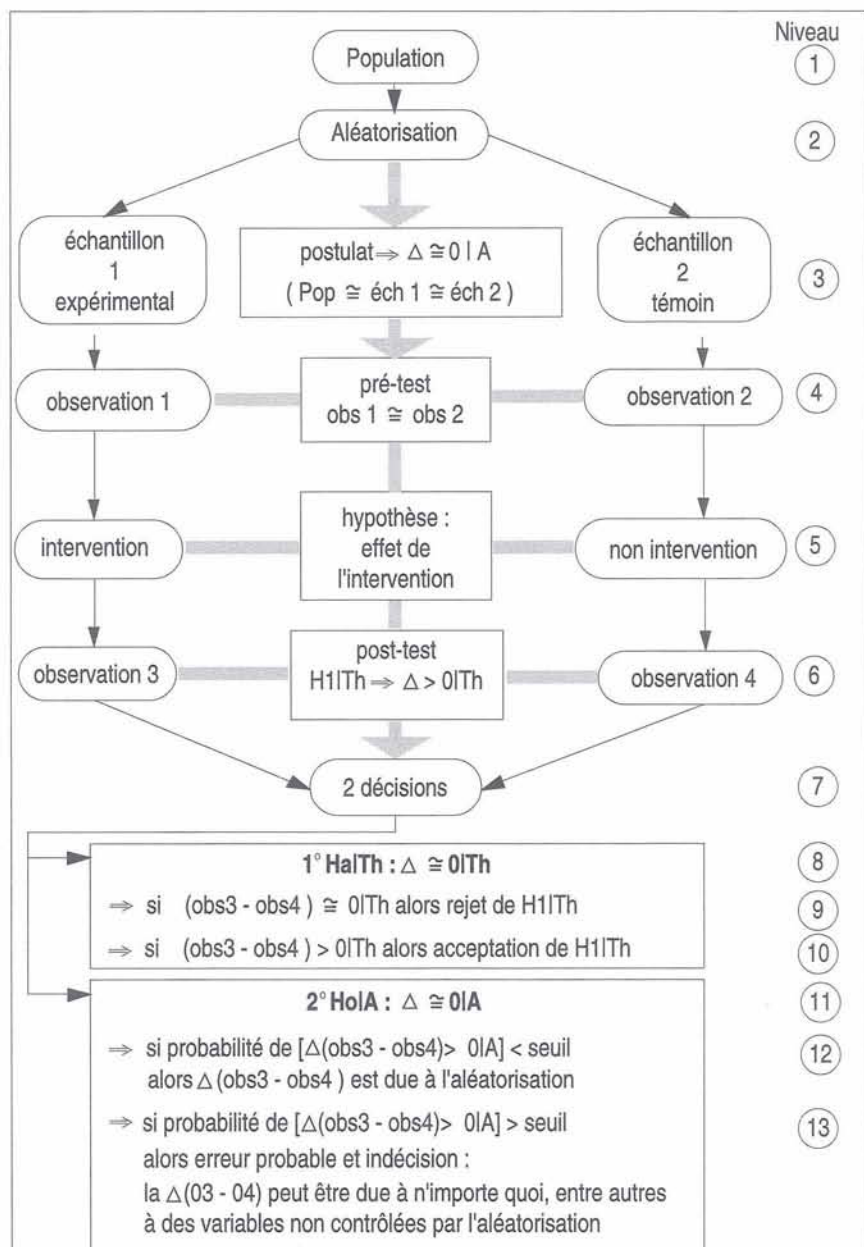


Figure 8.5 : L'organigramme des processus et des décisions dans un plan expérimental avec deux groupes aléatorisés.

Or, comme on le constate dans de nombreuses recherches, cette décision sur l'hypothèse de travail ($H1|Th$) s'effectue souvent par la réfutation de son contraire, l'hypothèse alternative ($Ha|Th$, niveau 11). Étant donné que, par faiblesse de la théorie, les chercheurs ne peuvent prédire exactement l'effet attendu sous l'hypothèse de travail, ils formulent une hypothèse alternative annonçant qu'il n'y aura pas de différence entre les observations après l'intervention. L'hypothèse alternative par rapport à la théorie (intervention), soit $Ha|Th \Rightarrow \Delta (O3 - O4) \approx 0|Th$, se formule d'une manière semblable à l'hypothèse nulle postulée grâce à l'aléatorisation, soit $Ho|A \Rightarrow \Delta (O3 - O4) \approx 0|A$, mais elles n'ont pas la même signification. L'observation d'une différence nulle entre $O3$ et $O4$ conduit au rejet de l'hypothèse de travail ($H1|Th$), quand la décision porte sur l'effet de l'intervention, alors que l'observation d'une différence statistiquement non significative ($\approx 0|A$ avec $p > \text{seuil}$) conduit au rejet de l'hypothèse nulle ($Ho|A$, selon laquelle la différence est due aux aléas) et à une indécision : on ne sait pas trancher si la différence observée sous une telle probabilité est due à l'intervention (la variable indépendante) ou à d'autres variables présentes lors de l'échantillonnage et non contrôlées par celui-ci. C'est seulement lorsque la probabilité associée est significative ($\approx 0|A$ avec $p < \text{seuil}$) que l'on admet que la différence est seulement due à l'aléatorisation. Dans ce cas, le hasard ayant un effet nul, la différence observée est probablement dépendante des éléments du plan de la recherche. Mais si l'on peut rejeter un effet d'échantillonnage aléatoire, cela ne veut pas dire automatiquement que l'effet observé est dû à la variable indépendante ou à l'intervention : l'effet pourrait dépendre d'autres variables, comme les variables parasites et les variables non contrôlées par le plan de recherche.

4.3.3 Le contrôle des variables parasites

Lorsqu'il apparaît que la variation observée de la variable dépendante est suffisante pour penser à une dépendance de la variable indépendante (dans la confirmation), ou lorsque la variation ne semble pas suffisante pour justifier une dépendance (dans l'infirmité), le chercheur ne peut pas conclure avant d'avoir effectué, comme dans une recherche exploratoire, la vérification des comptes. Si la décision sur la signification des résultats dépend des seuils et des critères préalablement retenus, il faut examiner le déroulement de la recherche afin d'y dépister tout ce qui pourrait affaiblir la démonstration. Dans le cas d'une recherche comportant des échantillons aléatoires, la possibilité d'un biais d'échantillonnage par le hasard constitue la première variable parasite à vérifier et à cette fin, mais à cette fin

seulement, le test statistique est indispensable. Mais il faut aussi vérifier l'impact des autres variables parasites qui auraient pu interférer avec la relation mise à l'épreuve. Ces variables parasites (voir chapitre 9) sont tous les facteurs autres que la ou les variables indépendantes, qui pourraient expliquer la relation observée. Une liste sommaire de ces parasites possibles comporte : le passé et la maturation des sujets, des événements extérieurs concomitants à la recherche, des effets de sélection, l'effet des instructions qui sont données et des consignes implicites à la situation de recherche, la désirabilité sociale, etc.

Ce sera seulement lorsque la possibilité d'un impact de ces variables aura été éliminée qu'une décision pourra être prise sur la signification des résultats. La possibilité de l'interférence d'une des variables parasites restreint la certitude de la décision. Or, un seul plan de recherche permet de tester une relation entre variables sans ambiguïté et hors de tout doute raisonnable, c'est-à-dire en contrôlant l'interférence des variables parasites. Malheureusement, la possibilité de le réaliser, en éducation, est une exception tellement rarissime qu'il faut le considérer comme un idéal utopique : c'est le plan expérimental avec aléatorisation de l'échantillonnage et des conditions, comportant au moins six groupes dont deux groupes avec placebo (groupes à anticipation), deux groupes sans observation initiale, et mise en place d'une procédure de double aveugle au niveau des observations et de l'intervention ou de la manipulation (voir chapitre 10).

4.3.4 Les incertitudes de la décision dans la vérification confirmative

Pour finir, il faut noter deux faiblesses de la vérification confirmative.

La première faiblesse est inhérente aux plans utilisés. Étant donné que l'hypothèse est une application particulière (une opérationnalisation) déduite d'une thèse abstraite (à portée générale), la vérification de la thèse ne porte que sur un échantillon des situations qui la concernent et non pas sur la population de ces situations. On ne vérifie la thèse que par la confirmation d'une hypothèse, alors que plusieurs hypothèses auraient pu être déduites de la même thèse. La confirmation est donc inductive et la thèse est seulement plausible puisqu'on n'a pas fait la preuve que toutes les situations concernées s'y soumettent. De ce point de vue, la déduction de la recherche infirmative est logiquement plus forte, puisque tant qu'on n'a pas trouvé une infirmation, la thèse est tenue pour valide.

La seconde faiblesse provient de la précipitation avec laquelle on entreprend fréquemment de telles recherches. Sans avoir eu le temps

d'élaborer une théorie (une chaîne de relations) suffisante pour construire des instruments de mesure précis et pour prédire l'ampleur, le sens et la fréquence des variations à observer, la vérification porte plus souvent sur des intuitions que sur des hypothèses. Par manque de fondement théorique, les hypothèses sont réduites à prédire qu'il y aura un changement, mais sans être capables de préciser lequel. De ce fait, les chercheurs se contentent de la simple prédiction d'une différence quelconque entre les situations et ils se sentent confirmés dès que la comparaison entre les groupes donne une différence non nulle⁶. Accepterions-nous sans sourciller que les météorologues se contentent de prédire que le temps ne sera pas le même demain qu'aujourd'hui ? Pourtant c'est exactement ce que font de nombreuses recherches tant en sciences de l'éducation que dans les autres sciences humaines et dans les sciences biomédicales. Trop souvent on se contente de prédire que l'effet de l'intervention se manifesterá par une différence de performance entre les groupes, sans pouvoir ou sans risquer d'annoncer l'ampleur, la direction, la fréquence et la généralité (tous les sujets ou certains types de sujets) de la différence attendue. Les tristes mésaventures des «mystères de la mémoire de l'eau⁷» sont là pour nous consoler; mais ce n'est pas parce que nous ne serions pas les seuls à manifester de telles faiblesses que nous ne devons pas tenter de les corriger.

La solution aux faiblesses de la recherche vérificative réside dans une multiplication patiente de recherches exploratoires, dans l'élaboration de théories plutôt que dans la vérification d'intuitions et dans la construction d'outils de mesure assez précis, qu'ils soient quantitatifs ou qualitatifs, pour mettre à l'épreuve des hypothèses rationnellement déduites plutôt que des paris. Tant que les théories seront faibles, les mesures seront grossières et ceux qui disent faire de la recherche seront plus des parieurs que des chercheurs.

4.4 L'étude de cas vérificative

La comparaison qui a été faite des deux tactiques infirmatives suggère que les études de cas peuvent aussi être utilisées pour vérifier une théorie. La démarche suit alors celle de l'expérience critique : il suffit qu'un cas ou

6 C'est sans doute parce qu'ils testent leur hypothèse de recherche (une différence due au traitement) par l'épreuve de son contraire (une différence nulle due au traitement) qu'ils ont assimilé l'épreuve de leur hypothèse au test statistique de l'hypothèse d'une différence nulle entre échantillons aléatoires. Mais ce sont bien deux choses différentes.

7 de Pracontal M., *Les mystères de la mémoire de l'eau*. Paris : Éditions La Découverte, 1990.

qu'un élément d'un cas ne corresponde pas au modèle théorique pour que la théorie soit remise en cause. Dans cette perspective vérificative, on cherchera non pas ce qui ressemble au modèle théorique, mais plutôt ce qui, dans le cas, s'en écarte afin de pointer quels éléments de la théorie doivent être revus et corrigés. Une telle procédure a été utilisée afin de vérifier un modèle de l'enseignement déduit de la théorie cybernétique et mathématique de la communication. Après avoir enregistré une micro-leçon portant sur une notion, la chercheuse⁸ effectue une analyse conceptuelle de la notion enseignée pour en construire le réseau conceptuel. Ensuite, elle applique le modèle théorique de la communication à ce réseau pour élaborer le modèle théorique de l'enseignement de cette notion. À partir de ce modèle, elle prédit quels seront les effets sur l'apprentissage de la présence ou de l'absence des différents segments de l'enseignement s'il est effectué selon le modèle théorique. Elle compare ensuite le modèle théorique de l'enseignement de la notion avec la micro-leçon qui a été effectivement donnée. Elle constate alors que le modèle de l'enseignement déduit de la théorie de la communication était bien pertinent pour expliquer un certain nombre de faits observés, mais qu'il était limité aux aspects cognitifs et devait être complété sur différents aspects, dont le climat social de la classe. La figure 8.6 illustre une telle exploitation vérificative de l'étude de cas.

L'utilisation de l'étude de cas unique a donc un intérêt majeur tant dans une perspective d'exploration que dans une perspective de mise à l'épreuve d'hypothèse. On remarquera cependant qu'étant donné la lourdeur de cette dernière utilisation, la démarche n'en vaut la peine que si la théorie que l'on veut tester ou améliorer est déjà suffisamment riche pour permettre des prédictions précises et donc pour être confrontée à une expérience critique. Les chercheurs ne devraient se lancer dans la vérification d'un énoncé théorique que s'ils peuvent prédire avec précision quels seront les indices de la présence de l'effet et quels seront les indices de son absence.

8 Huot J., *Éléments pour un modèle de la communication pédagogique*. Thèse de doctorat en sciences de l'éducation. Faculté des sciences de l'éducation, Université de Montréal, 1979.

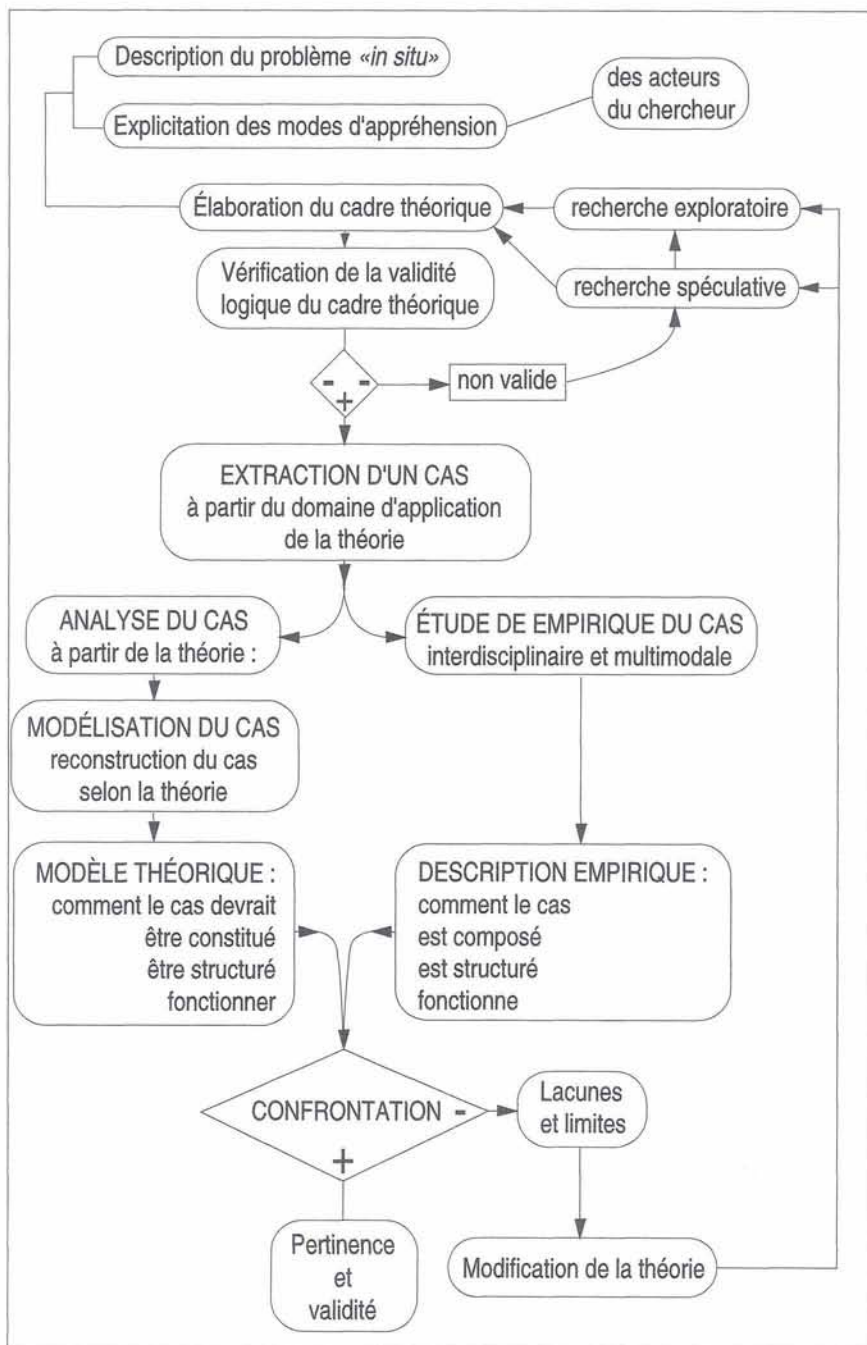


Figure 8.6 : L'étude de cas comme démarche vérificative.

LES PLANS DE RECHERCHE

- 1 Du désordre apparent à l'ordre grâce à l'organisation
- 2 Les plans de recherche à organisation temporelle
 - 2.1 *Les études longitudinales*
 - 2.2 *Les études rétrospectives*
 - 2.3 *Les études transversales*
 - 2.4 *Les séries temporelles*
- 3 Les plans de recherche à organisation spatiale
 - 3.1 *L'échantillonnage et l'aléatorisation*
 - 3.2 *Les variables parasites et les plans expérimentaux*
 - 3.3 *Les plans quasi expérimentaux, appariés et a posteriori*
- 4 Les plans mixtes à organisation spatio-temporelle
- 5 Les plans à plusieurs dimensions : plans factoriels et carrés
- 6 Les études de cas simples et croisées

1 Du désordre apparent à l'ordre grâce à l'organisation

Les techniques de constitution des données, envisagées au chapitre 4 et qui seront détaillées dans la troisième partie (chapitres 13, 14 et 15), sont la plupart du temps insuffisantes à la formulation ou à la vérification d'hypothèses. Dans la mesure où l'objectif de toute recherche consiste à mettre de l'ordre dans le désordre apparent (observer des structures émergentes, identifier des relations), la meilleure stratégie ne résidera pas dans la contemplation du matériel recueilli. Fivaz¹ a montré que le chercheur doit être actif dès le départ s'il veut se donner quelques chances de trouver quelque chose. Dans la plupart des cas, on ne peut percevoir quelque chose de nouveau que si on change de point de vue ou si les éléments à percevoir changent de position. Même les animaux le savent : la proie qui ne veut pas être repérée par un prédateur s'immobilise. Pour qu'une structure puisse apparaître, le chercheur suscite une organisation en imposant un essai d'ordre à ce qui apparaît désordonné. Si l'organisation suscitée est pertinente, une structure apparaîtra; sinon le matériel restera illisible.

Puisque la manière la plus efficace pour identifier une structure consiste à comparer un ensemble non structuré avec un autre qui l'est, le recueil des données s'effectuera selon une organisation qui s'appuie sur le temps ou sur l'espace (qui sont les structures de base de notre existence). Le chercheur pourra alors confronter le matériel recueilli avec la structure choisie pour le constituer. En plus de faciliter l'émergence des relations que l'on cherche, l'organisation spatio-temporelle de la constitution des données inhibe l'interférence de facteurs ou de variables que l'on veut contrôler si elle est rationnellement construite. La construction d'un plan spatio-temporel de constitution des données permet donc de susciter l'apparition de structures et d'argumenter, par le contrôle des variables parasites, les interprétations qui seront proposées. Enfin, les manières de traiter les données ne sont pas indifférentes au plan de leur constitution, et certains types de traitement réclament que les données aient été constituées selon un plan qui respecte des exigences particulières.

2 Les plans de recherche à organisation temporelle

Une première organisation de la constitution des données conduit à l'élaboration de plans temporels parce que, dans le domaine de l'éducation, le facteur temps joue généralement un rôle important. D'une part, le temps

1 Fivaz R., *L'ordre et la volupté*. Lausanne : Presses polytechniques romandes, 1989.

permet le développement et la maturation des processus et des capacités qui évoluent avec l'âge. D'autre part, la durée permet l'instauration d'un apprentissage ou son raffinement en habiletés par des exercices et des répétitions diverses. C'est aussi la durée qui instaure l'oubli dont l'observation permet de mesurer la réussite ou l'échec de l'apprentissage. Selon que l'étude porte sur un phénomène qui évolue plus ou moins naturellement ou sur un changement provoqué par une intervention, on parlera d'étude longitudinale et transversale ou de recherche par série (ou séquence) temporelle.

2.1 Les études longitudinales

Quelle que soit la technique de constitution des données (observations, entrevues, enquêtes, sondages, tests, etc.), les études longitudinales consistent à recueillir des données auprès d'un même sujet ou d'un même groupe de sujets à plusieurs reprises plus ou moins espacées dans le temps et sur une période relativement longue. Ainsi dans certains travaux, comme ceux de Gesell en psychologie génétique, les chercheurs ont pris des mesures et effectué des observations tous les ans sur la même population de sujets de façon à pouvoir décrire la psychologie des enfants de 1, 2, 3, 4, et 5 ans et de montrer quelle en est l'évolution. Ce type de recherche implique de suivre les sujets et elle rencontre deux exigences importantes : contrôler la perte de sujets (qui déménagent, abandonnent l'école, etc.), ce qu'on appelle la «mortalité expérimentale», et adapter les techniques de mesure et d'observation aux nouveaux comportements qui apparaissent avec le temps. Mais il s'agit de la meilleure manière d'observer une évolution.

L'étude longitudinale présente une difficulté supplémentaire dans le cas de recherches qui doivent être menées dans des délais restreints (par exemple, pour l'obtention d'un diplôme) ou qui portent sur des phénomènes à évolution lente ou dont l'évolution n'est perceptible qu'à la suite d'une longue et patiente observation : la disparition du chercheur. Il n'y a pas que des chercheurs qui perdent leurs sujets, il y a aussi des sujets qui perdent leurs chercheurs !

2.2 Les études rétrospectives

Afin de pallier cette difficulté, deux stratégies ont été développées, chacune avec leurs difficultés et leurs lacunes. Puisqu'on n'a pas le temps de voir les individus évoluer et qu'on ne peut pas leur demander de dire à l'avance ce qu'ils deviendront, la première stratégie consiste à faire une

recherche rétrospective. Ne pouvant observer ce qui met trop longtemps à se développer, la constitution des données s'appuie sur la mémoire humaine pour essayer de reconstruire, après coup, ce qui s'est produit. Ainsi procèdent des méthodes qui portent sur le témoignage du passé, comme les histoires de vie, les récits de pratique et les enquêtes rétrospectives qui espèrent pouvoir reconstruire l'évolution des choses en remontant d'abord dans le temps, puis en rétablissant le fil de l'histoire à partir du souvenir de l'origine. Mais on devine les failles de cette stratégie : le souvenir fait souvent défaut, il ne garde que les événements qui ont eu une charge affective ou cognitive importante, il a tendance à exagérer les émotions attachées aux événements, à les dichotomiser et à privilégier le positif ou le négatif selon les humeurs du moment. De plus, le souvenir reconstruit le passé en rendant le récit plus harmonieux, plus cohérent et plus rationnel que les événements ont pu l'être. Aussi la recherche rétrospective ne peut pas se contenter d'une seule source d'information : à côté du souvenir de l'acteur, il faut chercher les documents d'archives et le témoignage de compagnons qui donnent des assises aux données recueillies et qui permettent, par recoupement, de pondérer la valeur du récit reconstruit.

2.3 Les études transversales

La seconde solution de remplacement à l'étude longitudinale consiste à faire une étude transversale dans laquelle on observe, au même moment, différents groupes d'âges différents du point de vue du phénomène que l'on veut étudier. Comme ces études comparent différents groupes, on pourrait les classer parmi les recherches à organisation spatiale. Mais le critère de sélection des groupes est bien temporel : l'âge. Plus précisément, l'âge est pas nécessairement biologique ni psychologique dans ces études, il est en rapport avec l'événement à observer. Lorsqu'il s'agit de l'observation d'attitudes qui dépendent de l'âge biologique, celui-ci servira de base : on étudiera simultanément des groupes d'enfants de 1, 2, 3, 4, 5 et 6 ans de manière à voir, par exemple, comment se construit progressivement la coopération avant l'entrée à l'école primaire. Dans d'autres études, le critère sera la durée de l'exposition à une situation donnée : par exemple, sans égard à leur âge biologique, on comparera les représentations de l'informatique que s'en font des novices qui la découvrent, des habitués qui ont terminé leur apprentissage, et des experts. Dans d'autres cas, l'étude portera sur des générations successives, comme dans les études sur l'image de la femme que se font des jeunes adolescentes, leurs mères, leurs grands-mères et, lorsqu'elles vivent encore, leurs arrière-grands-mères. Les défauts

de ces recherches sont aussi évidents. Chacun des groupes d'une étude transversale ne diffère pas seulement sur la variable temps, il diffère aussi sur la variable histoire et sur la variable culture. Non seulement leur expérience du phénomène étudié est plus ou moins longue, mais leur passé aussi est différent : ils n'ont pas vécu plus de choses, ils ont vécu d'autres choses. Il faudra essayer de démêler ce qui est dû à la durée de l'expérience et ce qui est dû à l'histoire. On peut tenter de dégager les particularités propres au passé des différents groupes d'âge en combinant l'étude rétrospective sur les groupes plus âgés avec l'étude transversale.

2.4 Les séries temporelles

Lorsque la recherche porte sur l'effet d'une intervention dans le temps, la stratégie générale est celle de la série temporelle. Il s'agit d'enregistrer les conduites du ou des sujets avant l'intervention, afin de connaître leur *ligne de base*, puis de poursuivre les enregistrements pendant et après l'intervention afin de noter si un effet se développe, se maintient ou se réduit. Le bloc S.1 du tableau 9.1 illustre une telle série temporelle.

S.1	temps :	1	2	3	4	5	6	7	8	9	10	...	n
	intervention :					A							
	observations :	1	2	3	4	5	6	7	8	9	10	...	n
S.2	temps :	1	2	3	4	5	6	7	8	9	10	...	n
	interventions :					A		A		A			
	observations :	1	2	3	4	5	6	7	8	9	10	...	n
S.3	temps :	1	2	3	4	5	6	7	8	9	10	...	n
	interventions :					A1			A2			A3	
	observations :	1	2	3	4	5	6	7	8	9	10	...	n
S.4	temps :	1	2	3	4	5	6	7	8	9	10	...	n
	interventions :					A	A	A	A	A			
							B	B	B	B	B		
								C	C	C	C	C	
	observations :	1	2	3	4	5	6	7	8	9	10	...	n

Tableau 9.1 : Exemples de séries temporelles : avec une intervention en S1, avec la même intervention répétée en S2, avec une variation dans l'intervention en S3, avec une intervention multiple, maintenue et décalée dans S4. Dans chacun de ces exemples, les observations faites aux temps 1 à 4 servent à établir la ligne de base.

Plusieurs variantes de la série temporelle ont été développées sous l'appellation de plans à cas unique². En effet, une fois la ligne de base établie par plusieurs observations successives, on peut, après une intervention initiale, arrêter l'intervention puis la reprendre une deuxième, une troisième, une énième fois (comme en S.2 dans le tableau 9.1), ou encore la maintenir un certain temps. Dans certaines études pharmacologiques ou psychologiques, le plan pourra être encore plus sophistiqué : par exemple, on pourra progressivement augmenter les doses du traitement (drogue en pharmacologie, renforcement ou autre système de stimulation ou d'inhibition en psychologie) ou faire varier le type de traitement (voir le schéma du bloc S.3 dans le tableau 9.1). On peut trouver des études plus complexes dans lesquelles plusieurs traitements se rapportant à des types différents de comportement sont entrepris en phases décalées dans le temps (comme schématisé en S.4 dans le tableau 9.1). Les diverses variantes de la série temporelle sont aussi nombreuses que l'imagination des chercheurs est riche.

Ces plans de recherche par série temporelle posent évidemment quelques problèmes, chaque nouvelle variante essayant de corriger les défauts des précédentes mais apportant aussi ses propres failles. Une discussion détaillée des problèmes posés lors de l'analyse des données que ces plans produisent, se trouve dans des ouvrages spécialisés cités dans le livre de Robert² ou dans l'ouvrage classique de Cook et Campbell³.

3 Les plans de recherche à organisation spatiale

Dans les plans à organisation spatiale, la constitution des données s'effectue de telle sorte que l'on peut comparer des groupes d'individus, supposément semblables, soumis au même moment à des conditions différentes. L'objectif n'est pas d'observer une évolution, mais bien d'identifier les différences dues aux conditions auxquelles les groupes sont soumis.

Le plan de base se résume à la simple comparaison de deux descriptions; par exemple, la comparaison des performances obtenues à un même test de français par les élèves de la fin du primaire selon qu'ils

2 Voir par exemple A. Fortin et M. Robert, Plans de recherche à cas unique, chapitre 7, dans M. Robert (éd.), *Fondements et étapes de la recherche scientifique en psychologie*. Montréal : Chenelière et Stanké, 1982, p. 133-151.

3 Cook T.D. et D.T. Campbell, *Quasi-experimentation. Design & Analysis Issues for Field Settings*. Chicago : Rand McNally, 1979.

proviennent de l'école publique ou du secteur privé. Mais un tel plan est rarement praticable, car l'examen de toute une population est vite coûteux et lourd. Aussi, pour se faire une idée des performances de cette population (les élèves de la fin du primaire) sous les deux conditions (école publique et école privée), on utilise des échantillons d'élèves tirés de cette population sous chacune des conditions (ce qui constitue alors deux sous-populations) et l'on fait une inférence : on prend la performance manifestée par les échantillons comme une estimation des performances qui seraient observables dans les populations si leur observation pouvait se faire. Le tableau 9.2 schématise le plan de base à partir duquel de telles observations sont construites.

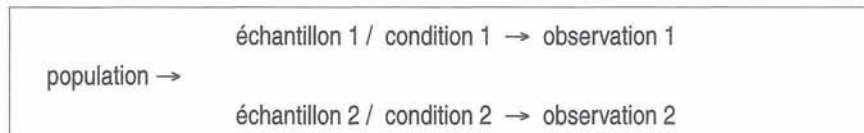


Tableau 9.2 : Plan spatial élémentaire : la comparaison des observations de deux échantillons d'une même population sous deux conditions.

On note immédiatement que les différences entre les observations 1 et 2 peuvent provenir d'autres sources que les différences entre les conditions 1 et 2. Par exemple, il est évident que les différences de performance (observations 1 et 2) entre les élèves des écoles publiques et privées (conditions 1 et 2) pourraient provenir d'autres facteurs, comme l'appartenance sociale et culturelle : les deux échantillons pourraient déjà être différents, indépendamment de leur fréquentation d'une école publique ou privée. Dès lors, ce plan de base est insatisfaisant et doit subir l'une ou l'autre de deux corrections possibles, soit

- élargir l'analyse des conditions à l'ensemble des facteurs qui surdéterminent la performance au test et ainsi transformer ce plan en une étude de cas croisée (voir plus loin en 9.5);
- introduire une organisation plus sophistiquée de l'expérience et en faire un plan expérimental, quasi expérimental ou à groupes appariés.

3.1 L'échantillonnage et l'aléatorisation

Si l'on veut pouvoir observer l'effet potentiel d'une intervention (recherche exploratoire) ou si l'on veut mettre à l'épreuve une hypothèse en propos de l'effet d'une intervention (recherche vérificative), la manière la moins contestable de procéder consiste à montrer, par un prétest, que deux

groupes de sujets sont identiques avant l'intervention mais qu'au post-test, après l'intervention, seul le groupe qui a subi l'intervention manifeste un changement (exploration) ou présente le comportement attendu (vérification). Une telle comparaison, dite expérimentale, implique deux conditions : que l'on puisse introduire de manière contrôlée l'intervention dont on veut étudier l'impact, et que les deux groupes de sujets soient identiques sur l'ensemble des facteurs qui pourraient avoir une influence sur les résultats au post-test ou dont l'influence pourrait interférer avec l'intervention.

La première condition, soit l'introduction contrôlée de l'intervention, a des conséquences directes sur la situation étudiée : les expérimentations exigent souvent des situations artificielles, elles se réalisent idéalement en laboratoire, et lorsque l'on veut les réaliser sur le terrain, elles impliquent souvent des modifications à l'organisation «naturelle» de l'environnement. Or ces modifications, qui réduisent la validité écologique de l'expérience, sont souvent difficiles à réaliser en milieu scolaire pour des raisons éthiques autant qu'administratives et syndicales.

Il arrive que l'intervention ait été ou soit introduite indépendamment du chercheur, et cela d'une manière contrôlée, du moins à première vue. C'est ce qui se produit lorsque l'intervention découle d'une décision administrative ou politique appliquée à une partie du système d'éducation; il est alors possible d'en observer les effets potentiels en comparant la partie du système qui a subi la décision avec la partie du système qui ne l'a pas subie. Mais, dans ces cas, le chercheur ne peut se contenter d'observer que la décision a été prise; il doit s'assurer qu'elle a bien été mise en application, ce qui impose souvent une intrusion dans le système. De plus, hormis de rares exceptions, ce genre de recherche est le plus souvent rétrospective : les événements se sont déjà produits lorsqu'on décide de les étudier. On retrouve dans ce cas un plan temporel rétrospectif, qui peut s'enrichir pour devenir une étude de cas, que l'on peut ensuite comparer (croiser) avec d'autres études de cas.

La deuxième condition est pratiquement impossible à réaliser, sinon en recourant à un artifice statistique sur la base du postulat déjà mentionné au chapitre 8. En effet, pour constituer deux groupes identiques, il faudrait d'abord établir l'inventaire exhaustif des facteurs ou des traits susceptibles de produire des effets concurrents à l'intervention. Or, la possibilité d'un tel inventaire est douteuse sur le plan théorique, car l'exhaustivité d'un tel inventaire implique que l'on connaisse suffisamment les conditions et les effets de l'intervention, et si justement on les connaissait déjà, la recherche

ne serait plus à faire. Ensuite, en admettant que l'impossible soit réalisé, il faudrait examiner (ou tester) la possession de l'ensemble des traits de cet inventaire sur des sujets en un nombre suffisant pour constituer deux groupes de sujets semblables par rapport à chacun des éléments de cet inventaire. Or, voilà une deuxième exigence impossible : l'on sait maintenant que même au point de vue biologique, il n'y a pas deux individus semblables.

La solution réside donc dans un artifice statistique sur base d'un postulat. On suppose d'abord que l'on dispose de l'inventaire des traits qui pourraient, hypothétiquement, interférer avec l'intervention. Ensuite, on postule que les variations de la possession de ces traits se distribuent de manière aléatoire dans la population. Dès lors, si ces deux présuppositions sont acceptées, il suffirait de tirer au hasard des individus d'une population nombreuse et de répartir ces individus au hasard dans les groupes que l'on va comparer, pour poser l'hypothèse d'une différence statistiquement nulle entre ces groupes. En fait, par *l'aléatorisation*, c'est-à-dire en utilisant une extraction au hasard de la population et en faisant l'allocation des sujets au hasard, on construit des groupes qui ne sont pas identiques mais qui sont probablement équivalents. Leur similitude est approximative, la probabilité d'une différence étant simplement due aux aléas de l'échantillonnage (revoir le paragraphe 4.3.2 du chapitre 8) .

Une fois les groupes extraits et constitués au hasard à partir d'une population, le hasard devrait encore intervenir une troisième fois : dans la répartition des conditions expérimentales sur les groupes constitués. Ce triple recours au hasard, qui constitue au sens strict une *aléatorisation* (ou «randomisation» en français), garantit la puissance des plans expérimentaux, car c'est sur cette base que les effets de la sélection, ou des différences de passé des sujets, peuvent être contrôlés.

Rappelons que le test statistique ne vérifie pas du tout l'effet possible de l'intervention; il ne fait que renseigner sur la possibilité d'un biais, d'une variable parasite due à l'échantillonnage au hasard rendu nécessaire pour constituer des groupes équivalents. Or, ce biais est probable, étant donné qu'il n'est pas du tout certain que le postulat à propos de la distribution des traits dans la population soit aussi robuste qu'on le souhaite. D'où l'importance d'attribuer son sens exact au test statistique, de lui faire dire ce qu'il dit et seulement ce qu'il peut dire.

3.2 Les variables parasites et les plans expérimentaux

Les plans expérimentaux ont été développés de manière à contrôler les variables parasites, dites aussi alternatives (à ne pas confondre avec l'hypothèse alternative, voir chapitre 8), dont les effets peuvent interférer avec l'intervention pour en augmenter les effets, pour les masquer, les neutraliser, ou même pour produire des effets inverses à ceux recherchés.

Les principales variables parasites sont les suivantes :

1° *L'histoire* : un événement concomitant à l'intervention interfère avec elle et est responsable de tout ou d'une partie de l'effet attribué erronément à l'intervention.

2° *La maturation* : l'évolution spontanée des sujets interfère avec l'intervention et l'on attribue erronément à l'intervention ce qui n'est qu'un effet naturel de développement des sujets.

3° *La sélection* : des traits, des caractéristiques propres aux différents (groupes de) sujets les rendent plus ou moins sensibles à l'intervention, si bien que l'on peut attribuer erronément à celle-ci un effet qui serait dû, partiellement ou totalement, au passé des sujets.

4° *La mortalité* : à la suite de diverses causes, la disparition de certains sujets, par exemple des sujets peu sensibles, crée un effet dont l'ampleur est injustement attribuée au traitement.

5° *La non-standardisation* : les conditions d'observation ou de mesure ne restant pas identiques au prétest et au post-test. En conséquence, les résultats changent entre les deux observations à cause d'une modification des conditions d'observation et non pas à cause de l'intervention.

6° *L'apprentissage du testing* : les sujets, à l'affût d'indices leur permettant de bien faire, apprennent au prétest ce qu'ils doivent faire au post-test et leurs résultats sont dus à cet apprentissage plutôt qu'à l'intervention. On appelle aussi ce phénomène la *réactivité* de la mesure.

7° *La régression statistique* : les performances de sujets qui, pour divers facteurs dont le stress, produisent des résultats accidentellement extrêmes au prétest (très faibles ou très forts), redeviennent normales en cours d'expérience et produisent des résultats au post-test qui convergent vers la moyenne. Cette convergence interfère avec les mesures de l'effet dû à l'intervention.

8° *La stimulation expérimentale* : le fait de participer à une expérience crée une stimulation des sujets, et cette stimulation produit un effet semblable à celui recherché par l'intervention. Cette variable est aussi appelée *effet Hawthorne*.

9° *Les consignes implicites* : le matériel de l'expérience, la disposition et l'organisation des lieux, les attitudes des chercheurs constituent des indices qui permettent aux sujets de deviner ce qu'ils doivent faire.

10° *L'observation sélective des cas favorables* : les chercheurs connaissant mieux les traces des phénomènes qu'ils cherchent que les traces de la multitude des phénomènes qu'ils ne cherchent pas, ont tendance à surtout percevoir les événements favorables à l'intervention.

11° *L'effet d'échantillonnage au hasard* : la différence qui subsiste entre les groupes à la suite de l'échantillonnage au hasard (l'aléatorisation), crée une différence dans les données dont la probabilité excède ce que les conventions de recherche du domaine tolèrent, cette différence étant injustement attribuée à l'intervention.

Les premiers plans expérimentaux ont été imaginés pour contrôler l'effet des variables parasites : ou bien il s'agissait de limiter leur impact, ou bien de l'identifier et de l'isoler de l'effet de la variable indépendante sur la variable dépendante. Tel qu'illustré au tableau 9.3, ces plans comportent deux groupes obtenus par aléatorisation : un groupe expérimental auquel on applique l'intervention et un groupe contrôle (ou témoin) qui ne bénéficie pas de l'intervention.

	Gr. Exp.	Obs.1	interv.	Obs.2
Population → aléatorisation				
	Gr. Ctr.	Obs.3	rien ⁴	Obs.4

Tableau 9.3 : Plan expérimental à deux groupes.

On perçoit que l'organisation de la constitution des données, selon un tel plan, permet de contrôler les cinq premières variables parasites. Les diverses comparaisons que l'on peut faire entre les quatre observations permettent d'évaluer l'effet relatif de ces variables et de l'intervention. Par exemple, si l'effet de la maturation sur les sujets du groupe contrôle (Obs. 4 > Obs. 3) se marque autant sur ceux du groupe expérimental, on peut départager l'effet potentiel du traitement par la comparaison de l'écart

4 *Rien* veut dire ici seulement absence de l'intervention dont on veut étudier les effets; la plupart du temps cela signifie que l'on continue à s'occuper des sujets comme d'habitude.

sur le groupe contrôle avec celui observé sur le groupe expérimental (Obs. 2 - Obs. 1). Si au prétest les groupes sont équivalents (Obs. 1 \equiv Obs. 3), la différence entre l'écart observé sur le groupe contrôle, attribuable à la maturation, et l'écart observé sur le groupe expérimental (soit : [(Obs. 2 - Obs. 1) - (Obs. 4 - Obs. 3)]) donne la part d'effet attribuable au traitement.

Campbell⁵ a montré cependant que pour contrôler l'effet des variables comme la réactivité de la mesure, la régression statistique et la stimulation expérimentale, il faut ajouter au plan de recherche deux groupes qui ne subissent pas la mesure initiale, ce qui constitue un plan à quatre groupes, dit de Solomon. Il se schématise comme dans le tableau 9.4. Dans ce dernier, étant donné que les groupes Exp.2 et Ctr.2 sont équivalents à Exp.1 et Ctr.1 grâce à l'aléatorisation, l'écart entre les observations initiales effectuées sur ces derniers groupes (Obs.1 et Obs.3) avec les observations finales effectuées sur les groupes sans prétest (Obs.5 et Obs.6) fournit une estimation de l'effet de l'intervention [(Obs.1-Obs.5)-(Obs.3-Obs.6)], tout en contrôlant la réactivité de la mesure et la stimulation expérimentale qui sont absentes dans Exp.2 et Ctr.2.

	Gr.Exp.1	Obs.1	interv.	Obs.2
	Gr.Ctr.1	Obs.3	rien	Obs.4
Population → aléatorisation				
	Gr.Exp.2	rien	interv.	Obs.5
	Gr.Ctr.2	rien	rien	Obs.6

Tableau 9.4 : Plan expérimental à quatre groupes ou plan de Solomon.

Mais nous montrons aussi au chapitre 10 que les variables parasites qui composent les *effets Rosenthal et Pygmalion*, imposent d'ajouter encore au moins deux autres groupes dits à *anticipation*, comme dans le tableau 9.5. Ces deux groupes supplémentaires, soit un groupe expérimental à anticipation et un groupe contrôle à anticipation qui reçoivent chacun une pseudo-intervention (ou placebo), permettent de comparer l'effet de l'intervention réelle (Obs.2. et Obs 7.) avec l'effet dû aux attentes (Obs.6 et Obs.9) que les sujets et les chercheurs peuvent projeter dans l'expérience par des consignes implicites et par l'observation sélective.

5 Campbell D.T., Facteurs intéressant la validité des études expérimentales dans des contextes sociaux. Dans Lemaine G et J.M. Lemaine, *Psychologie expérimentale et expérimentation*, Paris : Mouton-Bordas, 1969, p. 47-61. Pour une analyse plus fouillée et comportant les aspects statistiques, voir Campbell D.T. et J.C. Stanley, *Experimental and Quasi-experimental Design*. Dans le *Handbook for Research on Teaching*. Chicago : Rand McNally, 1966.

Population → aléatorisation	Gr.Exp.1	Obs.1	interv.	Obs.2
	Gr.Ctr.1	Obs.3	rien	Obs.4
	Gr.Ant.1	Obs.5	placebo	Obs.6
	Gr.Exp.2	rien	interv.	Obs.7
	Gr.Ctr.2	rien	rien	Obs.8
	Gr.Ant.2	rien	placebo	Obs.9

Tableau 9.5 : Plan expérimental à six groupes dont deux à anticipation.

Encore faut-il noter qu'un tel plan à six groupes n'est efficace que si on y applique la *technique du double aveugle* ou méthode en *double insu*. Dans cette technique, tous les acteurs sont aveugles : les sujets ne doivent pas savoir à quel groupe ils appartiennent, les observateurs ne doivent pas savoir quel type de sujet ils observent. De plus, ceux qui s'occupent des sujets (les enseignants, mais parfois même la direction de l'école) ne doivent pas savoir si ce qu'ils font participe à l'expérience, s'il s'agit du traitement (intervention) ou du pseudo-traitement (intervention aux effets neutralisés). Un plan avec anticipation (comme celui du tableau 9.5) et double aveugle, soit une expérience en double aveugle appliquée aux six groupes, comporte évidemment un ensemble d'exigences très difficiles à respecter dans les situations de recherche en éducation, alors que seul ce dernier type de plan permettrait une véritable mise à l'épreuve de l'effet de l'intervention. Aussi, ce plan devient un idéal auquel on se réfère pour estimer quels sont les contrôles auxquels on a renoncé, et donc quelles sont les incertitudes qui ont été introduites, lorsqu'on choisit un plan moins parfait mais applicable.

3.3 Les plans quasi expérimentaux, appariés et a posteriori

Comme déjà noté, les recherches se déroulant sur le terrain ne peuvent que difficilement rencontrer les exigences du *double aveugle* et celles de *l'aléatorisation*. Lorsque cette dernière est sacrifiée, les observations initiales produisent des résultats différents de groupe à groupe. On dit de ces plans qu'ils sont *quasi expérimentaux*. Ils supposent des analyses plus complexes (appelées de *covariance*) lors de l'évaluation de l'effet de l'intervention au post-test, pour tenir compte de l'écart qui existait entre les performances initiales au prétest.

Le problème de l'équivalence initiale des groupes peut cependant recevoir une solution intéressante par la méthode des plans avec *groupes appariés*. Cette solution implique que les principaux traits pertinents à l'intervention peuvent être définis sur une base théorique et faire l'objet de mesures dont on postule la validité pour les fins de la recherche. Cette solution est aussi tout indiquée lorsqu'on ne dispose pas d'une population assez nombreuse pour en extraire des échantillons, mais qu'au contraire on peut travailler sur la population en entier, car il s'agit de population comportant des traits assez rares. La technique des groupes appariés consiste à diviser la population totale en grappes (ou sous-ensembles de sujets qui se ressemblent entre eux) composées d'autant de sujets semblables que de groupes prévus au plan de recherche. Comme illustré dans la figure 9.1, on commence par un prétest de classement dans lequel on mesure, chez tous les sujets, les traits jugés pertinents, y compris la variable dépendante (prétest sur le comportement qui fournira un indice de l'effet potentiel). Ces mesures faites, on dresse le profil de chacun des sujets, c'est-à-dire son portrait composé de l'ensemble des traits mesurés. Ensuite on calcule un indice de similitude entre les profils (par exemple, le nombre de traits partagés divisé par l'ensemble des traits), puis on constitue les sous-ensembles (paires, triades, quatuors, etc.) appelés *grappes*, composés chacun d'autant de sujets que le plan d'expérience comporte de groupes. Le plus souvent, on tente de minimiser les différences entre les membres d'une même grappe et à maximiser les différences entre les grappes. Le nombre de grappes correspondra au rapport entre le nombre d'individus dans la population et le nombre de sujets qui composent une grappe. Des techniques informatisées permettent de constituer de telles classifications (procédures de classification automatique ou d'analyse en grappes - dites *clustering* et *clusters analysis* en anglais). Une fois ces grappes composées, on répartit chacun de leurs membres dans les groupes correspondant aux conditions de l'expérience. De la sorte, les groupes ne sont pas équivalents, mais chaque membre de chaque groupe a son semblable (un membre de sa grappe) dans les autres groupes.

Une fois le post-test effectué, on peut ne plus se contenter de comparer entre eux les résultats moyens de chacun des groupes comme dans les plans expérimentaux ou quasi expérimentaux. On peut, en outre, comparer ce que sont devenus les individus provenant d'une même grappe. Cela enrichit l'analyse : on ne se limite pas à dégager des effets moyens, on peut aussi dégager les effets différentiels qui seraient fonction des traits typiques à chacune des grappes. Ce type d'analyse fournit souvent des renseignements

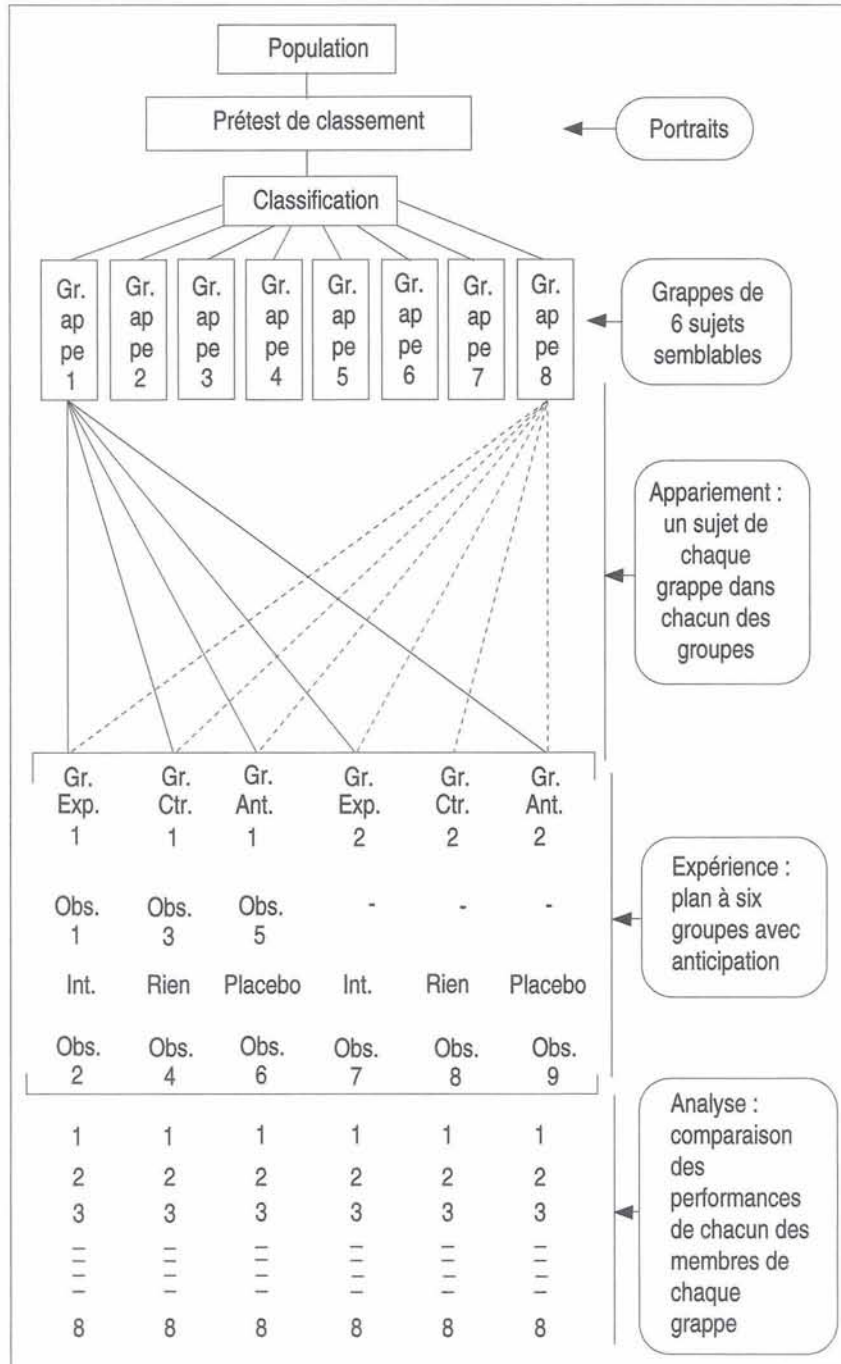


Figure 9.1 : Plan avec six groupes appariés et anticipation.

impossibles à obtenir lorsqu'on compare des moyennes de groupe, car dans les groupes on peut trouver des sujets qui réagissent positivement à l'expérience et d'autres qui y réagissent négativement, les résultats de ceux-ci annulant les résultats de ceux-là, ce qui produit des moyennes insignifiantes. Il n'est pas rare qu'en éducation un procédé didactique, par exemple, soit efficace avec une portion de la classe et pas avec une autre. Les plans comportant un appariement des groupes permettent, plus que les autres, d'accéder à cette information.

Les techniques d'analyse des résultats habituellement utilisées dans les plans expérimentaux et quasi expérimentaux suivent un raisonnement inférentiel de type prédictif : si le traitement ou l'intervention a bien les effets que la théorie lui attribue ou que l'intuition anticipe, alors les résultats au post-test doivent témoigner de l'effet hypothétiquement prédit à partir de la théorie ou attendu par l'intuition. Il suffit donc de comparer les résultats au post-test des groupes expérimentaux avec les résultats attendus (si la théorie a permis une telle prédiction) et avec les résultats des groupes contrôles et des groupes à anticipation qui ne devraient pas montrer l'effet hypothétiquement prédit pour le groupe expérimental. Cependant, ce raisonnement assez robuste ne tient que lorsque les interventions ont des effets assez simples et que l'on effectue des observations sur des situations et avec du matériel dont on connaît et dont on contrôle tous les composants. Or, les situations de recherche en éducation sont souvent complexes et engendrent des phénomènes surdéterminés, c'est-à-dire des phénomènes dont l'apparition ou les modifications dépendent de la conjonction de plusieurs facteurs dont on ne peut déterminer à l'avance quelles combinaisons seront nécessaires et suffisantes à la production des phénomènes en question. Dans le cas de tels phénomènes, la logique des plans expérimentaux semble impropre à la découverte de nouvelles relations.

De façon à mieux respecter la logique de la découverte plutôt que celle de la vérification, et en tenant compte de la surdétermination des phénomènes éducatifs, nous avons proposé et utilisé une autre stratégie d'analyse des résultats : *la comparaison des performances a posteriori*⁶. Cette procédure, illustrée à la figure 9.2, comporte aussi la constitution de groupes expérimentaux, contrôle et à anticipation comme dans les plans classiques ou dans les plans avec groupes appariés. Mais les résultats au post-test ne sont pas comparés entre eux immédiatement, c'est-à-dire que

6 Van der Maren J.M., Revoir la recherche en éducation : cesser de prédire pour mieux comprendre. *Repères, essais en éducation*, Faculté des sciences de l'éducation, Université de Montréal, 1986, n° 6, p. 100-140.

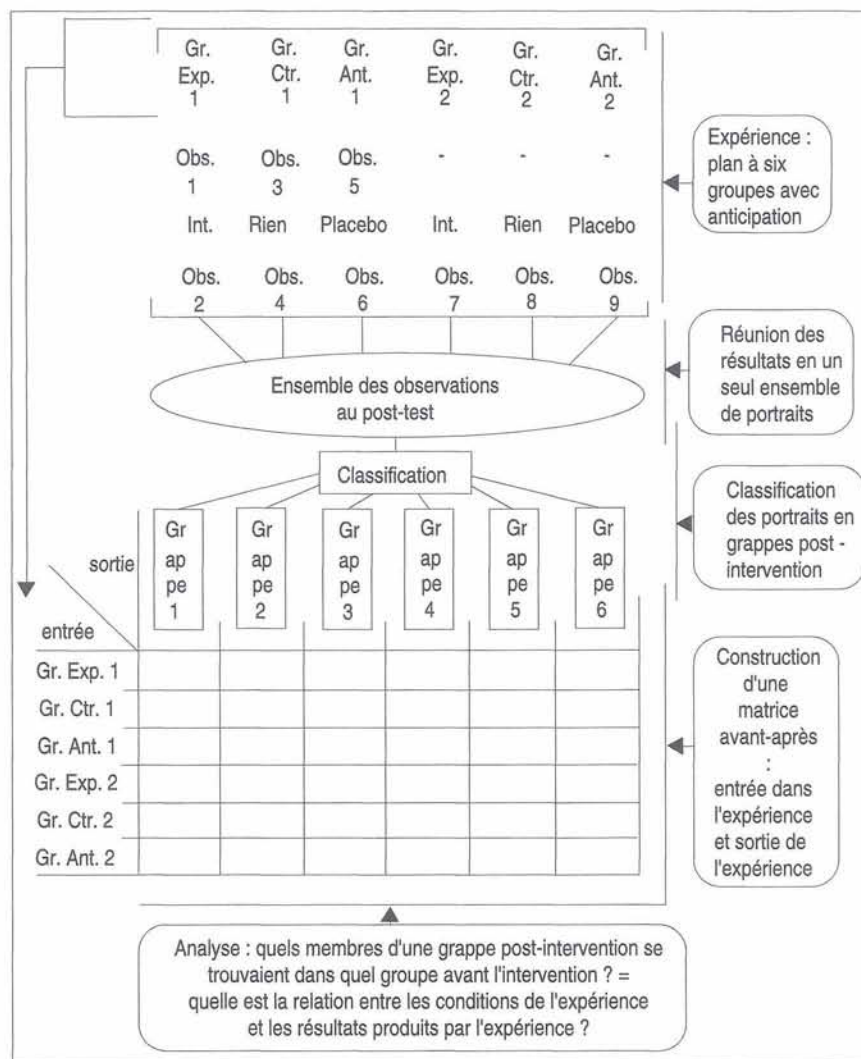


Figure 9.2 : Comparaison des performances *a posteriori* pour un plan à six groupes avec anticipation

les calculs ne se font pas sur la base des groupes de sujets issus des conditions de l'expérience. On commence d'abord par mettre ensemble tous les résultats du post-test et on applique sur cet ensemble de comportements post-intervention une classification en grappes selon une procédure semblable à celle suivie pour constituer des groupes appariés. Cette classification produit donc un certain nombre de grappes de sujets qui sont rassemblés non pas sur la base des conditions d'expérience qu'ils ont subies,

mais bien sur la base de la similitude des mesures ou des observations de leurs comportements après l'intervention. Habituellement, on tente de construire autant de grappes post-intervention qu'il y avait de groupes dans le plan d'expérience, afin de pouvoir composer une matrice carrée. On dispose à ce moment de deux classements des sujets : le classement selon les conditions d'expérience appliquées aux groupes (c.-à-d. le classement *a priori*, ou sur l'entrée de l'expérience) et le classement selon les résultats au post-test (c.-à-d. le classement *a posteriori* ou sur la sortie de l'expérience).

La relation entre les conditions d'expérience [(intervention, pseudo-intervention, absence d'intervention) x (avec et sans prétest)] et les résultats au post-test est mise en évidence en composant une matrice ou table à double entrée. Dans cette matrice, les rangées correspondent aux groupes affectés aux conditions d'expérience (entrée) et les colonnes aux grappes construites sur les résultats au post-test (sortie); chacune des cellules au croisement des colonnes et des rangées comporte le nombre de sujets ayant fait partie à la fois de tel groupe *a priori* (sur les conditions) et de telle grappe *a posteriori* (sur les résultats). On peut ensuite calculer différents coefficients sur de telles matrices, comme des coefficients de contingence ou comme le coefficient «*l*» (logarithmique) de Spitz⁷, transformables, enfin, en coefficients de détermination. Ces derniers coefficients indiquent la partie de l'effet observé sur les résultats qui est attribuable aux conditions de l'expérience, ce qui correspond à une logique plus adaptée aux caractéristiques et aux contraintes de la recherche exploratoire en éducation.

4 Les plans mixtes à organisation spatio-temporelle

La nécessité de mieux mettre en évidence l'effet et le maintien de l'effet d'une intervention auprès de différents types de sujets et dans des conditions différentes, a suscité des plans combinant les séries temporelles avec les plans expérimentaux et quasi expérimentaux. Ainsi, comme l'illustre le tableau 9.6, plusieurs groupes (organisation spatiale) peuvent être observés à plusieurs reprises espacées dans le temps avant et après une intervention. Dans un tel plan, la durée des effets (variable dépendante) après l'intervention (variable indépendante) est mise en relation avec une ou plusieurs caractéristiques qui peuvent interférer (variable modératrice) avec la variable

7 Delvaux J.P., Calcul simplifié de la corrélation entre deux variables. Estimation rapide du coefficient de corrélation de Bravais-Pearson à partir d'un test de contingence entre deux séries de données paires. *Revue de psychologie et des sciences de l'éducation*, 1970, 5, p. 302-308.

indépendante. Le tableau 9.6 pourrait illustrer une recherche visant à découvrir la durée de l'effet d'une variable indépendante (un changement de l'évaluation, par exemple) sur l'attention en classe (variable dépendante observée), étant donné des groupes d'écoliers pris à quatre niveaux scolaires différents (variable modératrice). Les résultats des recherches faites selon de tels plans s'expriment aisément par des graphiques comportant plusieurs courbes.

Temps :	1	2	3	4	5	6	7	8	9	10	...	n
Intervention :					A							
Observations Gr.Exp.1 :	1	2	3	4	5	6	7	8	9	10		n
Observations Gr.Exp.2 :	1	2	3	4	5	6	7	8	9	10		n
Observations Gr.Exp.3 :	1	2	3	4	5	6	7	8	9	10		n
Observations Gr.Exp.4 :	1	2	3	4	5	6	7	8	9	10		n

Tableau 9.6 : Plan mixte avec quatre groupes exprimant une variable modératrice.

Dans une autre variante de série temporelle, plusieurs variables dépendantes sont observées simultanément chez un même groupe avant et après l'intervention. Par exemple, comme dans le tableau 9.7, on peut observer l'évolution des performances scolaires, du climat social de la classe et de la motivation des élèves (variables dépendantes) après une modification (variable indépendante) de la longueur des leçons et de la répartition des matières sur l'horaire quotidien et hebdomadaire.

Temps :	1	2	3	4	5	6	7	8	9	10	...	n
Intervention :					A							
Obs. performances :	1	2	3	4	5	6	7	8	9	10		n
Obs. climat social :	1	2	3	4	5	6	7	8	9	10		n
Obs. motivation :	1	2	3	4	5	6	7	8	9	10		n

Tableau 9.7 : Plan mixte avec observation de trois variables dépendantes.

Une forme assez intéressante pour l'observation de la persistance de l'effet d'une intervention consiste à constituer plusieurs groupes expérimentaux auxquels on applique, au même moment, une intervention dont les effets au post-test seront observés de manière décalée dans le temps pour chacun des groupes. On procède, par exemple, selon un plan semblable au tableau 9.8. Un tel plan a le grand avantage de permettre une observation

de la persistance (du temps 6 au temps n) de l'effet d'une intervention (appliquée au temps 5) en contrôlant l'interférence possible de l'observation sur la persistance. Ainsi, l'observation aux temps 7, 8 et 9 sur les groupes expérimentaux 2, 3 et 4, fournit une information sur l'effet sans que la persistance ait été stimulée par l'observation au temps 6 pour le groupe 2, aux temps 6 et 7 pour le groupe 3 et aux temps 6, 7 et 8 pour le groupe 4. Cette possibilité d'un post-test décalé qui n'est pas artificiellement stimulé par des observations intermédiaires, est essentielle lorsque les observations de l'effet (ou ses mesures) sont réactives (c.-à-d. produisent une interférence avec l'effet étudié).

Temps :	1	2	3	4	5	6	7	8	9	10	...	n
Intervention :					A							
Observations Gr.Exp.1 :	1	2	3	4	5	6	7	8	9	10		n
Observations Gr.Exp.2 :	1	2	3	4	5	-	6	7	8	10		n
Observations Gr.Exp.3 :	1	2	3	4	5	-	-	6	7	8		n
Observations Gr.Exp.4 :	1	2	3	4	5	-	-	-	6	7		n
Observations Gr.Ant.:	1	2	3	4	5	6	7	8	9	10		n
Observations Gr.Ctr.:	1	2	3	4	5	6	7	8	9	10		n

Tableau 9.8 : Plan mixte avec post-test décalé dans le temps sur les groupes expérimentaux.

Comme pour les autres types de plans, de nombreuses variantes s'inspirent d'un plan temporel initial, le modifient et le combinent avec un plan spatial. La prise en compte des variables alternatives doit se faire pour chacun de ces plans, en analysant les comparaisons possibles entre les conditions de chacun des groupes, leurs différentes lignes de base (observations antérieures à l'intervention) et les différences qui affectent les mesures aux post-tests.

5 Les plans à plusieurs dimensions : plans factoriels et carrés

Les phénomènes humains sont, on l'a dit, largement surdéterminés, c'est-à-dire qu'une variable peut dépendre de l'influence combinée de plusieurs autres variables; on définit d'ailleurs l'interaction comme l'influence de deux variables sur une troisième. Les *plans factoriels* ont été conçus pour essayer de mesurer l'effet conjoint de plusieurs variables sur

une autre, c'est-à-dire leur interaction. Théoriquement, un plan factoriel doit comporter autant de groupes qu'il y a de modalités (ou de niveaux) pour chacun des facteurs expérimentaux (ou variables indépendantes) en jeu dans la recherche. Cela veut dire que si le facteur 1 est représenté par «m» niveaux et le facteur 2 par «n» niveaux, le plan complet comportera «m fois n» situations expérimentales. Si l'on imagine un troisième facteur, il faudra encore multiplier le nombre de situations par le nombre de niveaux que ce facteur implique. La figure 9.3 illustre ce qu'exige un plan factoriel pour trois variables dont une variable dichotomique et deux variables à trois dimensions : 18 groupes.

On perçoit immédiatement que ce type de plan, idéal selon Matalon⁸, est lourd et coûteux. De plus, s'il permet aisément l'analyse du jeu de deux facteurs sur une variable dépendante, il devient difficilement interprétable dès que l'on ajoute un troisième facteur. Aussi, ce ne sera que dans de rares conditions de laboratoire que l'on pourra envisager ce plan, et ses prétentions formelles se trouvent fort limitées dans la pratique. Aussi nous nous contenterons de savoir que ces plans existent.

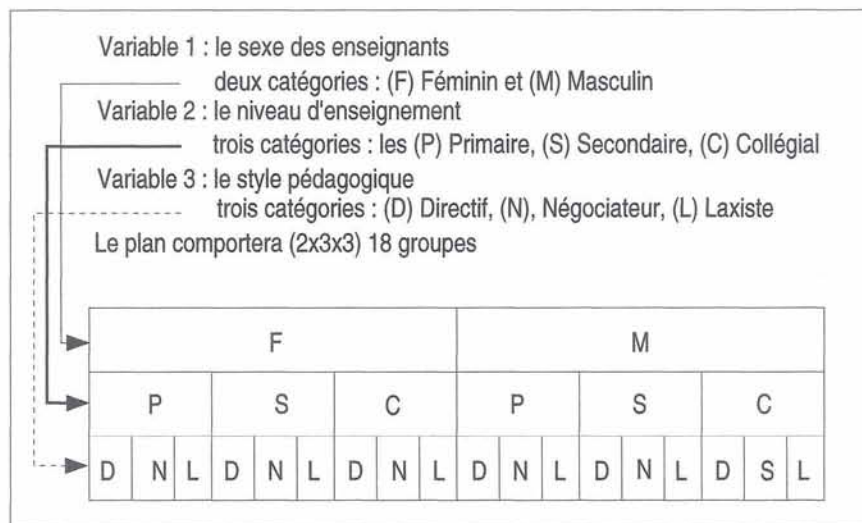


Figure 9.3 : Plan factoriel pour une recherche mettant en relation trois variables qualitatives.

8 Matalon B., La logique des plans d'expérience. Dans Lemaine G et J.-M. Lemaine, *Psychologie sociale et expérimentation*. Paris : Mouton-Bordas, 1969, pp. 33-45.

Mais il existe un plan factoriel partiel plus abordable, le plan en *carré latin*. Ce type de plan ne permet pas de vérifier les interactions d'ordre 2 (l'interaction des interactions, soit l'effet conjoint des interactions de plusieurs variables) et il doit comporter le même nombre de niveaux pour chacun des facteurs, ce que le plan factoriel n'exigeait pas. Plus léger que le plan factoriel, il nécessite une organisation plus souple tout en étant, lui aussi, consommateur de sujets. Par exemple, le tableau 9.9 montre que pour étudier l'effet de trois types de feed-back sur l'exécution d'une tâche, il faut construire un plan carré comportant trois groupes de sujets, trois tâches équivalentes appliquées aux temps 1, 2 et 3, et les trois types de feed-back. Mais l'ordre de succession des types de feed-back varie de groupe à groupe, de telle sorte que chaque groupe commence l'expérience par un feed-back différent.

	Temps 1	Temps 2	Temps 3
Groupe 1	FB 1	FB 2	FB 3
Groupe 2	FB 2	FB 3	FB 1
Groupe 3	FB 3	FB 1	FB 2

Tableau 9.9 : Plan contrebalancé pour une variable à trois dimensions (carré latin).

Une telle organisation de l'expérience, dite aussi en *plan contrebalancé*, permet d'annuler les effets de succession de l'ordre des feed-back dans la mesure où la rotation des conditions diffère de groupe à groupe. Comme on peut l'imaginer, des complexifications de ce plan existent. Dans l'exemple du tableau 9.10, on veut contrôler l'effet de la différence entre les tâches, même si elles sont considérées comme équivalentes, en plus des différences entre les feed-back. On a procédé en introduisant dans les cellules du plan illustré au tableau 9.9, une rotation supplémentaire des tâches, ce qui conduit à un plan dit en *carré gréco-latin* (les feed-back étant symbolisés par des caractères latins - A, B, C- et les tâches par des caractères grecs - α , β et γ).

	Temps 1	Temps 2	Temps 3
Groupe 1	FB 1 tâche 3	FB 2 tâche 1	FB 3 tâche 2
Groupe 2	FB 2 tâche 2	FB 3 tâche 3	FB 1 tâche 1
Groupe 3	FB 3 tâche 1	FB 1 tâche 2	FB 2 tâche 3

Tableau 9.10 : Plan contrebalancé pour deux variables à trois dimensions (gréco-latin).

6 Les études de cas simples et croisées

Les études de cas jouent, en éducation, trois rôles différents sous la même dénomination. L'étude de cas joue un rôle pédagogique comme méthode de formation : à partir du matériel extrait d'un cas, on peut procéder à des analyses et à des exercices de simulation. Mais ce rôle ne fait évidemment pas l'objet de ce manuel.

Dans la recherche, les études de cas peuvent tenir les deux positions situées aux extrêmes de la démarche hypothético-déductive. Elles ont été décrites au chapitre 8 en parlant de la monographie comme recherche exploratoire et des études de cas vérificatives. D'une part il s'agit d'une démarche inductive de compréhension, d'élaboration d'hypothèses : on prend une situation-cible, un cas que l'on circonscrit dans un espace-temps déterminé, et que l'on examine sous le plus grand nombre d'aspects possibles afin de le comprendre. D'autre part, il s'agit de confronter une théorie, ou le modèle qu'on en a déduit, avec la complexité d'une situation réelle, soit un cas situé dans son contexte; on examine alors quels sont les éléments du cas expliqués par la théorie ou le modèle, et quels en sont les éléments qui demandent de recourir à d'autres théories. La méthode des cas peut donc servir aux finalités inductives et déductives de la recherche : dans la perspective d'induire des hypothèses par lesquelles on tente de comprendre ce qui s'est passé, et dans celle de vérifier la pertinence d'une théorie existante en examinant comment elle peut expliquer ce qui s'est passé.

Leur caractéristique, en tant que plan de recherche, réside dans l'extension, sinon dans l'exhaustivité, du matériel recueilli sur une situation et dans l'interdisciplinarité des théories qui seront envisagées. Alors que les plans précédemment examinés utilisaient le temps et l'espace pour faire émerger ou pour vérifier une hypothèse, les études de cas recourent à la confrontation des points de vue. Ainsi, comme suggéré lors de l'analyse des plans de comparaison entre deux groupes (tableau 9.2), l'examen des effets d'une intervention, comme celui de la participation à deux réseaux différents d'éducation, ne peut se faire par une simple analyse des performances : pour arriver à comprendre les différences, il faut procéder à la description et à l'analyse de l'effet possible de l'ensemble des conditions sociales, culturelles, économiques, psychologiques, biologiques, et autres qui sont autant de caractéristiques qui peuvent intervenir dans le phénomène étudié. Autrement dit, les plans classiques envisagés jusqu'ici sont des plans de recherche centrés sur le jeu d'un nombre limité et d'une variété bien identifié de variables qu'ils tentent d'étudier sur un nombre assez élevé de

sujets. Par contre, dans les études de cas, on travaillera en n'examinant qu'un nombre réduit de sujets mais en envisageant le plus de variables possibles. L'étude de cas est par définition éclectique au sens où toutes les approches théoriques et toutes les méthodes de constitution de données qui peuvent fournir des éléments aidant à comprendre la complexité du phénomène, devront être envisagées.

Une étude de cas est dite simple lorsqu'elle porte sur un cas unique ou lorsque différents cas sont étudiés pour eux-mêmes, isolément les uns des autres. On parlera d'études de cas croisées (ou cas multiples) lorsque les catégories qui permettent l'analyse d'un cas sont comparées à celles qui sont utilisées dans d'autres cas. La comparaison des cas permet, non pas de généraliser, mais d'abstraire les éléments communs et d'identifier les particularités. Le croisement permet de valider les facteurs généraux propres à certaines conditions que l'on peut mettre en évidence par le recoupement des cas. En soi, cependant, le croisement de plusieurs cas ne constitue pas une forme d'évaluation de la fidélité ou de la transférabilité; il n'est pas équivalent à l'augmentation du nombre des sujets des échantillons d'une étude statistique. Il sert seulement de base à une démarche permettant d'abstraire des traits ou des facteurs communs qui fonderont l'induction d'une compréhension plus fine de ce qui est partagé par les différentes situations.

**LES BIAIS PSYCHOSOCIAUX
DE LA RECHERCHE
EN ÉDUCATION**

- 1 Introduction
- 2 Les contaminations dans le recueil des données
- 3 La contamination des résultats
 - 3.1 *Les contaminations dues aux élèves*
 - 3.2 *Les contaminations dues au personnel*
 - 3.3 *Les contaminations dues au chercheur*
- 4 La méthode du double aveugle
- 5 Le plan avec groupes à anticipation et double aveugle

1 Introduction¹

Parmi les efforts déployés pour donner un statut scientifique autonome aux recherches en sciences de l'éducation, on compte l'appropriation d'un certain nombre de techniques statistiques, de schémas expérimentaux, puis leur critique et leur adaptation aux conditions particulières de la discipline. Mais il semble que plus d'un chercheur résistent à admettre les conséquences de certains travaux méthodologiques qui, par ailleurs, ont ébranlé la conscience des praticiens. En effet, si dans certains cercles académiques, les travaux de Rosenthal, plus particulièrement le livre écrit en collaboration avec Jacobson², ont suscité maintes discussions, ils semblent n'avoir eu que peu d'effet sur l'élaboration des schémas de recherche. Ils ont, apparemment, été pris au pied de la lettre et embarqués au creux du débat anti-école avec le flux des analyses à propos de la fonction et du rôle de l'institution scolaire dans nos sociétés. Pourtant, les publications de Rosenthal avaient une autre visée : dévoiler les biais méthodologiques de la situation expérimentale dus aux attentes, aux anticipations des expérimentateurs. Si le psychologue social peut être suspecté d'influencer les performances de ses sujets par le fait même qu'il a des hypothèses, n'est-on pas en droit d'être encore plus suspicieux à l'égard du pédagogue ? N'est-il pas souvent plus engagé que le psychologue dans la défense de ses hypothèses ? Lorsqu'il s'agit d'évaluation de programmes, de méthodes ou d'outils pédagogiques, ne devrait-on pas manifester les mêmes exigences que lorsqu'il est question d'évaluer les effets d'une médication nouvelle ?

Certains projets de recherche ont attiré notre attention à la fois sur la bonne foi évidente et sur le souci de rigueur de pédagogues chercheurs et en même temps sur leur négligence de l'influence possible, et quasi certaine, de l'expérimentateur sur les résultats de ses travaux. D'ailleurs, ils sont bien souvent si sûrs de la qualité de leur démarche pédagogique qu'ils sont étonnés qu'on leur dise que ce ne sont là que des prophéties qui s'exaucent d'elles-mêmes. Cependant, les avertissements de Rosenthal étaient clairs, diffusés dès 1964, et présentés en français par Lemaine et Lemaine en 1969³.

1 Ce chapitre reprend l'essentiel d'un article publié : Van der Maren J.M., Le double aveugle contre Pygmalion : éléments de psychosociologie de la recherche en éducation et méthodologie des plans. *Revue des sciences de l'éducation*, 1977, III, 3, p. 365-380.

2 Rosenthal R. et L. Jacobson, *Pygmalion à l'école*. Traduction française à Tournai : Casterman, 1971.

3 Lemaine G. et J.M. Lemaine, *Psychologie sociale et expérimentation*, Paris : Mouton-Bordas, 1969, dont les chapitres suivants :

Matalon B., La logique des plans d'expérience, p. 35-45. .../...

L'examen d'un projet de recherche illustrera cette situation. Il s'agissait d'un travail visant à prouver l'effet sur le développement qualitatif d'une habileté académique (variable dépendante), d'exercices censés développer un présumé pré-requis (variable indépendante). Pour ce faire, les chercheurs prévoient un prétest et un post-test de cette habileté, l'entraînement au pré-requis prenant place entre les deux tests. De plus, ils décident de procéder selon un plan avec groupe expérimental et groupe contrôle. Mais, première remarque, ils ne prévoient pas de mesure pour ce pré-requis ! Or, il s'agit aussi d'une habileté qui se développe : certains de leurs élèves peuvent déjà l'avoir acquise, et son apprentissage peut résister aux exercices chez d'autres élèves. Ainsi, ils tentent de prouver l'existence d'un lien entre une variable dépendante (habileté) et une variable indépendante (pré-requis), sans évaluer les variations de cette dernière. D'une part, ce n'est pas parce que les chercheurs sont eux-mêmes responsables de l'enseignement de ce pré-requis qu'ils peuvent être sûrs de son effet; de plus, sans mesure, ils ne peuvent en connaître l'ampleur. Mais d'autre part, ils ne peuvent non plus avancer quoi que ce soit quant à la maîtrise de ce pré-requis dans le groupe contrôle (qui ne reçoit pas l'entraînement) tant au moment du prétest que du post-test. Dès lors, la vraie variable indépendante n'est plus le pré-requis, mais seulement la présence ou l'absence des chercheurs en tant qu'entraîneurs.

Ce premier défaut méthodologique, qui peut se corriger quand on s'en rend compte à temps, indique bien que l'implication du pédagogue chercheur dans sa pratique peut l'aveugler sur certaines conditions de sa recherche. Il attire aussi l'attention sur le risque d'une influence de l'implication du chercheur sur ses résultats. Poursuivant l'examen du projet, on constate que ce risque est bien grand. La recherche s'effectue simultanément dans deux milieux scolaires (aspect socio-économique). Dans chacune des écoles, il y aura groupe expérimental et groupe témoin, prétest, entraînement et post-test. Dans chaque école, le prétest et le post-test seront appliqués par le chercheur qui n'y est pas responsable de l'enseignement du pré-requis. De plus, les résultats au prétest ne sont pas communiqués à celui qui entraîne les élèves. Mais les groupes sont constitués *a priori* par les écoles (sur quelle base ?) et les épreuves seront appliquées en alternance, enfant témoin, enfant expérimental. En clair, cela signifie que le testeur,

.../...Campbell D.T., Facteurs intéressant la validité des études expérimentales dans des contextes sociaux, p. 47-61.

Rosenthal R., La participation volontaire, p. 71-79.

Riecken H.W., Psychologie sociale de l'expérimentation en psychologie : les consignes implicites et leurs conséquences, p. 283-290.

quel qu'il soit, saura, aux moments des pré- et post-tests, quel enfant recevra et aura reçu l'entraînement, ainsi que ceux qui ne le recevront pas. Il est clair, dès lors, qu'aucune garantie n'est présente contre un effet Pygmalion. Évidemment, le chercheur responsable de l'entraînement ignore les résultats au prétest, et ne peut donc modifier ses attitudes pédagogiques en fonction de ces résultats. Mais ici l'importance d'une telle modification est minimisée, puisqu'il n'a aucun contact avec les enfants du groupe témoin. Par contre, le testeur risque d'être plus qu'influencé : c'est lui qui aurait dû être dans l'ignorance de qui est membre du groupe expérimental ou du groupe témoin. En conséquence, dans un tel plan, en se partageant les rôles et les ignorances, les chercheurs jouent mal à l'aveugle : ils ne sont que borgnes et verront quand même les résultats influencés par leurs hypothèses. Car, même avec des épreuves ayant toutes les qualités psychométriques voulues, c'est dans une relation expérimentateur-sujet que s'effectuent les pré- et post-tests. Dès lors le testeur, connaissant l'appartenance des sujets à l'un ou l'autre des groupes, aura bien des difficultés à ne pas influencer les résultats de ceux-ci dans le sens de ses attentes. Les travaux de Rosenthal sont assez probants à ce sujet : même si des erreurs d'observation pourront, peut-être, être évitées, les réponses des sujets, de par les modes d'interaction qui se déploieront, seront plus que probablement influencées dans le sens des hypothèses mises à l'épreuve.

Ce cas en illustre bien d'autres et souligne la nécessité, pour les chercheurs en éducation, d'être attentifs à une série de contaminations⁴ possibles de leurs résultats, et de mettre en place des procédures du type *double aveugle* (ou double insu), des plans avec *groupe à anticipation* (avec pseudo-traitement) ou encore une combinaison des deux.

Quelles sont les contaminations possibles des observations ou des résultats dues aux attentes de l'expérimentateur ou même à la seule situation expérimentale ?

2 Les contaminations dans le recueil des données

Sans être exhaustif, soulignons quatre types de contaminations qui peuvent être assez fréquents dans le recueil des données en éducation.

Il y a d'abord l'effet de *stéréotypie*. L'observateur ou l'examineur classe, plus ou moins consciemment, le sujet dans une catégorie à laquelle

4 Bien que le terme «biais» soit classiquement utilisé pour recouvrir l'ensemble de ces problèmes, nous préférons «contaminations» qui n'a pas la connotation péjorative de «biais».

sont habituellement reliés certains traits. Ce premier classement peut s'être fait à partir d'une information préalable ou à partir de l'agglomération des premières perceptions. À la suite de quoi (et c'est là que se manifeste l'effet), l'examineur peut avoir tendance à ne plus percevoir, chez ce sujet, que les traits reliés à la catégorie dans laquelle le sujet a été classé, et à scotomiser les traits qui ne correspondraient pas au classement. Les examinateurs qui omettent de compter les erreurs dans les protocoles des bons élèves en sont un exemple fréquent.

Le classique effet de *halo* tend à s'en rapprocher. C'est la tendance à relier certains faits d'observation subséquents à une première série d'indices observés, sans qu'un lien objectif ne relie les deux séries d'éléments observés. Dans ce cas, le lien subjectivement établi entre les deux séries n'a d'autre fonction que de renforcer la première impression ou le premier jugement de l'observateur. Ainsi, un superviseur visitant un stagiaire qu'il estime empathique au bout de trois à cinq minutes, aura tendance, s'il n'y prend garde, à compter comme empathiques des comportements qui n'ont rien à voir avec ce concept rogérien. Toutes les bonnes attitudes seront mises au compte de l'empathie, et le stagiaire sera étiqueté comme vraiment empathique.

La propension que peut avoir l'observateur à noter plus ou moins d'événements en fonction de ses hypothèses est une autre contamination de l'observation. C'est l'*hyper- ou hypo-perception*. Il peut en aller de même lorsqu'il s'agit de dépouiller de façon interprétative des protocoles (par exemple, dans une analyse de contenu). Ces erreurs de perception, d'évaluation ou d'interprétation peuvent par ailleurs aller dans les deux sens : certains observateurs favorisant leurs hypothèses, d'autres au contraire les défavorisant dans un excès de rigueur.

Enfin, certains chercheurs en éducation pourraient avoir tendance à protéger leurs hypothèses en ne recueillant pas certaines catégories d'éléments ou de faits qui, bien que pertinents, risquent de rendre les résultats moins évidents. C'est un phénomène de *perception sélective*. Ainsi, il arrive que des chercheurs simplifient le champ théorique où s'insèrent leurs hypothèses et ne prennent en considération que certaines dimensions du phénomène étudié. Le mépris d'événements relevant d'effets appelés *secondaires* en constitue un exemple courant en éducation. L'effet peut être secondaire pour le chercheur, mais ne pas l'être pour la réalité scolaire. Dans certains cas, l'élimination de dimensions à observer provient de contraintes méthodologiques (lourdeur de l'instrumentation) mais, dans

d'autres cas, cela ressemble fort à la technique du vendeur qui, pour mieux vendre son produit, préfère en ignorer les défauts.

Soulignons encore une fois que ces contaminations de l'observation, tout comme les contaminations des résultats dont on parlera plus loin, peuvent être involontaires. Il s'agirait seulement, lorsqu'on les connaît, de prévoir des méthodes permettant sinon de les éliminer, au moins de les contrôler, de les mesurer.

3 La contamination des résultats

Les contaminations des résultats sont bien connues en psychologie sociale et peuvent tout autant se manifester en éducation. Envisageons d'abord les contaminations des résultats dues aux sujets. Le simple fait qu'ils se rendent compte de leur participation à une expérience va probablement modifier leurs comportements.

3.1 Les contaminations dues aux élèves

Bien que n'étant ni des volontaires ni des étudiants en psychologie qui tentent de se comporter en bons sujets (c'est-à-dire de manière à ce que les hypothèses de travail se trouvent confirmées), les élèves sont sensibles à toute expérimentation les concernant. Cela d'autant plus qu'il est bien rare qu'on puisse mener une quelconque recherche sur le terrain sans modifier l'environnement (classe, personnes, observateurs) ou la structure des conduites scolaires (horaires, manières de procéder, etc.). Dès lors, après un premier temps d'observation pendant lequel les élèves vont tenter de découvrir le sens de ce qui se passe autour d'eux, cette découverte va les amener à réagir de quatre manières différentes, ces réactions modifiant les résultats.

Un premier type de réactions se manifeste par une sorte d'*exaltation* à la suite de la découverte de leur importance, puisqu'on expérimente chez eux. Cette exaltation peut avoir des effets allant aussi bien dans le sens des hypothèses que contre elles, ou encore n'ayant rien à voir avec celles-ci, mais brouillant toute observation les concernant.

Au contraire, une forme d'*apathie* caractérise des élèves de certaines écoles, proches des centres universitaires ou cibles privilégiées des chercheurs étant donné leurs caractéristiques. Ces élèves en ont assez de voir leur environnement et leurs modes de fonctionnement sans cesse bousculés, et en arrivent à souhaiter ne plus être sujets d'expérience, afin

de bénéficier d'un enseignement normal (mais il y a beaucoup de chances pour qu'on ne demande pas leur avis!). Comme dans le cas de l'exaltation, les effets de cette attitude peuvent rendre les résultats hyper- ou hypo-concordants avec les hypothèses, ou encore tout neutraliser.

Ces deux premiers types de réactions produisent une surdité à des éléments des consignes de la recherche, car ce qui importe n'est pas le contenu de l'expérience, mais le fait qu'on expérimente. Ainsi, par exemple, dans une école pour déficients mentaux, tant au prétest et au post-test que pendant le traitement, les enfants avaient tendance à scotomiser l'aspect vitesse, de façon à rester plus longtemps en présence des expérimentateurs qui venaient de l'université, alors que cette tendance ne s'est pas manifestée lors d'une contre-expérience réalisée par l'intermédiaire du personnel habituel de l'établissement.

Deux autres types de réactions sont liés à la perception par les élèves d'indices qui prennent pour eux valeur de *consigne implicite* quant aux hypothèses en jeu ou quant à leur statut dans la recherche (par exemple, groupe expérimental vs groupe témoin). En ce qui concerne la perception d'indices quant aux hypothèses, la réaction des élèves dépendra essentiellement de leur position affective à l'égard des expérimentateurs; plus les élèves sont jeunes et plus ils sont socio-émotivement perturbés. Découvrant par les rumeurs qui circulent ou dans le déroulement de l'expérience, des indices quant au sens de celle-ci, ils pourront être tentés, ou bien de vouloir faire plaisir et dès lors de se comporter de manière à confirmer les hypothèses de départ (du moins ce qu'ils croient qu'elles sont), ou bien de faire enrager leur monde. Par cette dernière attitude, ils peuvent profiter de cette occasion pour régler le compte resté ouvert d'un conflit parfois ancien, en se comportant comme de mauvais sujets qui réagissent dans le sens opposé aux hypothèses perçues ou imaginées.

Enfin, la découverte d'indices quant à leur statut dans la recherche peut aussi amener les enfants à modifier leurs comportements et, en conséquence, à fausser les résultats. En effet, les élèves risquent de ne plus réagir de la même façon après la découverte de leur appartenance à un groupe témoin qui ne bénéficie pas du traitement offert au groupe expérimental, et inversement. Or, lorsque des modifications dans leur environnement ou dans les manières de les conduire interviennent, les élèves ont tendance à s'observer, à communiquer entre eux afin de savoir ce qui se passe; ainsi ils risquent de découvrir que le traitement n'est pas identique. Leurs réactions peuvent aussi être modifiées s'ils en viennent à croire qu'ils ont été sélectionnés en fonction de certains traits de personnalité, de certaines

de leurs caractéristiques académiques ou sociales. Lorsqu'enfin les élèves parviennent à combiner les deux séries d'indices (ceux sur les hypothèses et ceux sur leur appartenance aux groupes), leurs comportements manifesteront une tendance à s'aligner sur ce qu'ils croient être attendu d'eux ou, au contraire, à s'en écarter systématiquement. Quoique nous ayons rencontré un cas où la modification du comportement, bien que liée à la perception de ces indices, n'était pas fonction de la relation expérimentateur-sujet. Il s'agissait d'un adolescent classé caractériel, et plus particulièrement opposant, qui profita d'une expérimentation pour changer son attitude, ce qui annula un appariement; et cela parce qu'il en avait «assez de faire partie de la bande des caves».

3.2 Les contaminations dues au personnel

Si les élèves interprètent les éléments de la situation expérimentale en fonction de leur rapport à l'expérimentateur ou au chercheur, un risque semblable est encore plus grand lorsque la recherche porte sur la pratique pédagogique des maîtres ou des administrateurs. Lorsque ceux-ci acceptent de faire partie d'une recherche, qu'en espèrent-ils en retour ? Quels éléments de leur relation au chercheur ont pu déclencher leur participation à l'expérience ? En conséquence, quelles attitudes vont-ils avoir à l'égard des consignes ?

Dans certains cas, l'attitude est *a priori* positive et pourrait se rapprocher de l'exaltation que peuvent vivre les enfants. Participer à une telle expérience enflé les perceptions qu'ils se font d'eux-mêmes et les stimule parce que cela les rapproche de l'image de l'enseignant-chercheur, en éducation permanente, soucieux du progrès pédagogique. Ils adhèrent à une telle image idéale soit depuis leur formation, soit depuis leur rencontre avec certains animateurs ou encore depuis leur participation à l'un ou l'autre stage. Dès lors, la participation à une recherche satisfait leur narcissisme en restreignant l'écart entre l'image de soi et l'image idéale de soi pour soi, et les range dans une catégorie à part parmi les collègues : l'élite pédagogique des enseignants-chercheurs, des innovateurs, des progressistes.

D'autres, au contraire, risquent de réagir en fonction du dérangement que la recherche occasionne à leur routine. Leur participation ne sera pas très active, minimale au niveau de leur implication personnelle, même s'ils sont capables de présenter de grandes démonstrations d'intérêt, jusqu'au moment où, un élément de la recherche les ayant dérangés, ils réagiront négativement, sabotant même le travail entrepris. Ainsi, lors d'une recherche

projetée afin de vérifier certaines hypothèses quant aux phases de la négociation scolaire, l'administration d'une institution (avec internat) avait accepté l'ensemble du projet. Elle avait fourni du personnel (observateurs) pour deux mois, modifié les grilles horaires, consenti 2 250 \$ en matériel d'enregistrement et d'observation et entrepris des travaux dans une classe afin d'y installer un faux mur avec miroir sans tain. Une dernière contrainte avait été acceptée : la population de la classe où l'expérience se déroulerait devait rester stable. Or, à mi-chemin de l'expérimentation (après 15 jours), cette population fut changée aux 4/10^e parce que des lits étaient devenus inoccupés dans une section et que, pour les remplir, il fallait faire tourner les internes! Malgré l'investissement, l'expérience bien commencée ne put jamais être complétée.

Dans d'autres cas, l'attitude est plus mitigée : on n'y croit pas trop, on ne s'emballe pas, mais on va se comporter au mieux afin d'avoir une bonne note de la part de la hiérarchie, afin de se ménager une image aux yeux des collègues et des responsables de la recherche. Dès lors, si les premiers risquent, dans leur exaltation, de ne pas comprendre ce qui est attendu d'eux en faisant de la recherche d'un autre leur propre recherche (par exemple, modification des consignes pour adapter la tâche à leurs propres objectifs avec leur groupe d'enfants), ceux-ci vont tout faire pour satisfaire les attentes du chercheur. Entre autres, ils vont essayer de deviner ce qu'ils pourraient faire, au-delà des consignes qu'on leur donne, pour présenter l'image d'un bon collaborateur, qui fait ce qu'on lui demande et qui peut même en faire plus s'il croit que cela peut faire plaisir (par exemple, donner un entraînement spécial aux élèves en dehors des séances d'observation afin d'améliorer les performances, etc.).

Par ailleurs, certains percevront les phases de la recherche comme autant de formes d'inquisition. Ils y participent quand même pour ménager une image ou un statut, mais ils maquillent la réalité, non pas pour satisfaire les chercheurs, mais pour protéger l'image d'eux-mêmes ou de leurs élèves qu'ils croient devoir préserver. Ainsi, lors d'une pré-enquête afin de choisir des mesures d'habiletés professionnelles pour une recherche visant la validation d'une batterie d'orientation, un groupe d'enseignants avait systématiquement remonté les notes attribuées aux élèves. Cela parce que, pour eux, un bon enseignant était quelqu'un avec qui tous les élèves réussissent. Aussi, combien ne furent-ils pas dépités lorsque le chercheur confronta les notes attribuées sur les protocoles de tests avec les notes effectivement inscrites sur le bulletin remis aux élèves.

3.3 Les contaminations dues au chercheur

Si les résultats peuvent être contaminés par les enseignants, le chercheur peut, lui aussi, être responsable d'une semblable contamination; ce que Lemaine et Lemaine appellent l'effet Rosenthal, ce que d'autres ont appelé l'effet Pygmalion.

- | |
|---|
| <p>A. Contaminations dans le recueil ou l'interprétation des données</p> <p>A.1 Stéréotypie</p> <p>A.2 Halo</p> <p>A.3 Hyper-ou hypo-perception</p> <p>A.4 Perception sélective</p> <p>B. Contaminations dans la production des résultats</p> <p><i>B.1 Dues aux élèves</i></p> <p>B.1.1 Réactions au fait expérimental</p> <p>Exaltation</p> <p>Apathie</p> <p>B.1.2 Réactions aux données de la situation (consignes implicites)</p> <p>Par rapport aux hypothèses supposées</p> <p>Par rapport au statut du sujet (gr. exp. ou témoin)</p> <p><i>B.2 Dues aux maîtres et/ou aux administrateurs</i></p> <p>B.2.1 Réactions au fait expérimental</p> <p>Inflation de l'image et exaltation</p> <p>Défense contre l'évaluation</p> <p><i>B.3 Dues au chercheur : effet Rosenthal</i></p> <p><i>B.4 Dues à la conjonction de B.1, B.2 et B.3 : effet Pygmalion</i></p> <p>C. Contaminations et dans le recueil des données (A) et dans la production des résultats (B)</p> |
|---|

Tableau 10.1 : Les différents biais psychosociaux affectant la recherche en éducation.

Après avoir noté, dans de nombreuses anecdotes, de lourdes présomptions quant à ce phénomène, Rosenthal mit en place une série de recherches tendant à démontrer son existence. Les résultats qu'il rapporte et ceux d'autres chercheurs ne laissent pas de doute : dans bien des cas, les hypothèses que le chercheur tente de vérifier, contaminent ses résultats. Ses attentes influencent la performance des sujets, qui dès lors combinent ses espoirs. On ne peut, à première vue, empêcher un chercheur de vouloir

vérifier des hypothèses. Or, dès ce moment, il apparaît que ses attentes vont être communiquées aux sujets par des canaux variables dont le jeu n'est guère précisé à ce jour. Les sujets réagiront à cette communication inconsciente et involontaire comme ils réagissent à l'ensemble des consignes et des éléments significatifs de la situation. En conséquence, un certain nombre de recherches risquent de n'être que des prophéties qui se réalisent d'elles-mêmes.

Ce phénomène est d'autant plus probable que, comme Rosenberg (dans Rosenthal et Rosnow⁵) le signale, il semble exister une dynamique de la communication entre expérimentateur et sujet qui englobe plus que l'effet Rosenthal. Il y aurait une série d'interactions et d'ajustements entre les attentes de l'expérimentateur, la recherche par les sujets de consignes implicites (et conséquemment l'élaboration d'hypothèses quant à la conduite à tenir), les réactions de l'expérimentateur aux conduites des sujets, ces réactions étant ensuite perçues par les sujets comme des évaluations ou de nouveaux indices qui commandent un ajustement de leurs conduites. C'est une dynamique semblable qui règle normalement la communication entre maître et élèves dans les classes, chacun tentant d'ajuster sa conduite, son message à celui de l'autre. C'est là le véritable effet Pygmalion tel que l'illustre la pièce de B. Shaw.

Afin de mieux examiner l'ampleur du problème méthodologique posé et les stratégies possibles de contrôle, les principaux types de biais ou d'artefact jusqu'ici examinés ont été classés en trois catégories dans le tableau 10.1.

4 La méthode du double aveugle

Parmi l'inventaire des stratégies de contrôle des biais proposées ces vingt dernières années (voir, par exemple, Rosenthal et Rosnow), nous voulons insister sur la méthode dite du *double aveugle* ou encore du *double insu*, bien que, pour être correctement utilisée, elle exige souvent plus de deux «aveugles».

Le principe de base en est simple : les sujets ne doivent pas savoir s'ils font partie du groupe expérimental ou du groupe témoin, s'ils reçoivent ou non le traitement; l'observateur qui recueille les données tant au prétest, au post-test que pendant le traitement, ne doit pas savoir qui, parmi les sujets, fait partie du groupe expérimental ou du groupe témoin. De plus,

5 Rosenthal R. et R.L. Rosnow, *Artefact in behavioral*. New York Academic Press, 1969.

idéalement, l'expérimentateur qui applique le traitement ne doit pas savoir s'il applique le vrai ou le pseudo-traitement, ni quelles sont les hypothèses en question liées aux traitements. Aussi, comme nous allons l'illustrer à partir d'une étude pharmacologique réalisée en milieu orthopédagogique, cela aboutit souvent à faire beaucoup d'aveugles, chacun des acteurs engagés dans la recherche ignorant un petit quelque chose au sujet de l'ensemble de la recherche.

Afin de mieux saisir comment la méthode du double insu permet de neutraliser, sinon d'éliminer, certaines des contaminations envisagées, nous analyserons une recherche non publiée, n'ayant pas abouti aux résultats escomptés par le commanditaire. Il s'agissait d'une étude à propos des effets psychologiques d'un médicament «X» sur une population d'adolescents handicapés mentaux, mais ne présentant pas de perturbation neurologique décelable à l'électro-encéphalogramme et à l'examen clinique.

La stratégie de cette étude fut la suivante. Le superviseur du service psychologique de l'institution demande au psychologue d'appliquer aux adolescents une batterie de tests afin de préparer les dossiers de routine, comme il demande au psychiatre de lui indiquer les sujets présentant des perturbations neurologiques. Le psychologue applique les tests et transmet les résultats au superviseur. Celui-ci, sans avoir vu les adolescents, mais à partir des données fournies par le psychiatre et des résultats obtenus par le psychologue, constitue des paires de sujets en utilisant l'indice de similitude de Sorensen (tel que présenté par Dagnelie⁶). Mise au courant du nombre de paires composées par le superviseur, la firme lui expédie les flacons de pilules par paires, A et B, gardant pour elle le code de répartition du médicament et du placebo. Le superviseur étiquette les flacons avec les noms des adolescents selon les paires qu'il a composées et, ensuite, transmet les flacons à l'éducateur chargé de la distribution des médicaments. D'autres éducateurs (stagiaires) sont priés d'observer et d'annoter une série de comportements sur une liste sans qu'ils sachent à quelle partie de la paire appartient l'adolescent qui leur est assigné. À la fin de la période d'expérimentation, le psychologue applique à nouveaux les tests et en transmet les résultats au superviseur, qui reçoit des éducateurs les fiches d'observation complétées. Disposant de ces résultats, le superviseur établit ensuite les différences absolues pour chaque paire observée. Recevant seulement alors de la firme le code de distribution du médicament et du placebo dans les paires de flacon, il établit enfin le sens des différences et

6 Dagnelie P., Introduction aux problèmes et aux méthodes de classification numérique, *Biométrie-Praximétrie*, 1968, IX, 2, p. 87-III.

peut tester la signification statistique des changements intervenus (dans ce cas par un *Wilcoxon matched-pairs ranks test*).

Quelle est la position de chacun des intervenants dans cette expérimentation ?

- Le superviseur du service ne voit pas les adolescents. Il travaille uniquement sur des chiffres et, jusqu'au moment où il calcule le Wilcoxon, il ne sait pas à quelle série de données (partie des paires) correspond le traitement. Il sait quels types d'effets sont attendus du produit, ayant discuté avec la firme des éléments à retenir dans la liste d'observation et des tests psychologiques à appliquer.
- Le psychologue rencontre et applique aux adolescents une batterie de tests. Mais il ne sait qui est apparié avec qui (il n'est pas au courant du principe choisi pour établir les paires), et il ignore qui reçoit le médicament ou le placebo. Il sait seulement qu'une expérimentation est en cours mais il ne sait pas les effets attendus du produit. De plus, pour qu'il ne puisse pas les déduire, le superviseur ne lui a pas dit lesquels, parmi les tests de la batterie, ont été retenus pour la recherche.
- Les éducateurs observateurs, comme le responsable de la distribution des médicaments, sont dans la même situation que le psychologue. Les items de la liste d'observation couvrant un large éventail d'activités, il leur est aussi difficile d'en déduire les effets attendus.
- Les adolescents se rendent compte qu'on leur donne, à tous, une pilule supplémentaire. Peut-être certains s'interrogent-ils à ce sujet et remarquent que les éducateurs sont peut-être plus attentifs à leurs comportements. Mais le psychologue et les éducateurs ne sachant pas quels élèves appartiennent au groupe expérimental ou au groupe témoin et ne sachant quels sont les effets attendus (sinon une amélioration, mais sur quel plan ?), les indices à ce sujet sont semblables pour tous. Dès lors, si des modifications du comportement apparaissent à la suite de ces interactions, elles ne devraient pas être systématiquement orientées chez les uns plus que chez les autres.
- La firme pharmaceutique sait dans quels flacons sont le médicament et le placebo. Mais elle ne connaît ni la composition des paires, ni les individus qui reçoivent le produit actif, et elle n'est en contact ni avec les adolescents ni avec le psychologue et les éducateurs.

Ainsi, pendant cette expérimentation, ce n'étaient pas deux mais cinq groupes d'aveugles qui opéraient, chacun ignorant une partie de ce qu'il

aurait pu souhaiter savoir. Mais en utilisant un tel plan, nous n'avons pas pu éliminer toutes les occasions de contamination envisagées plus haut. Une majeure partie d'entre elles l'a cependant été. C'est ce que nous allons examiner avant d'envisager quel schéma idéal eut dû être mis en place.

Parmi les possibilités de contamination dans le recueil des données, celle à laquelle nous risquons d'avoir le plus donné prise, est la perception sélective (A.4). Les items de la grille d'observation et les tests psychologiques avaient été choisis en regard de certaines hypothèses : ils étaient censés pouvoir dépister les effets principaux du médicament. Les résultats ont conduit à conclure que, dans les conditions expérimentales données, aucun effet n'était apparu. Mais ceci ne vaut que pour les indices, les «comportements-signes», que nous avons choisis. D'autres tests et d'autres items n'auraient-ils pas révélé certains effets secondaires ? Secondaires pour nous à ce moment, mais peut-être importants par ailleurs.

Quant aux trois possibilités (A.1 à A.3), les effets en sont neutralisés, car étant donné la méthode utilisée, si elles s'appliquent, c'est à tous les éléments de la population, les témoins aussi bien que les expérimentaux. Par ailleurs, si dans ce cadre elles ont pu jouer, c'est plus en fonction de rapports interindividuels (entre adolescent et testeur ou entre adolescent et observateur) qu'en fonction de rapports entre chercheur (observateur ou testeur) et groupe de sujets (témoin ou expérimental).

Si l'on examine les possibilités de contamination des résultats (B), nous pouvons poser que la méthode des aveugles a permis d'échapper aux contaminations relevant des réactions aux données de la situation, que ce soit chez les sujets ou chez les expérimentateurs. De plus, l'effet Rosenthal (B.3.) y est neutralisé, s'il n'est pas absent : le manque absolu de contact entre le chercheur principal (la firme pharmaceutique et le superviseur) et les sujets minimise l'influence des hypothèses sur les résultats. Si cette influence peut passer par l'intermédiaire des observateurs ou du testeur, elle se voit neutralisée, puisque ces derniers ne savent pas à quel groupe appartient le sujet qu'ils ont devant eux. Ils ne peuvent ni lui coller une étiquette, ni s'inspirer de son histoire expérimentale, pour aligner leurs modes d'interaction avec lui.

Cependant, telle qu'utilisée, la méthode des aveugles ne permet pas d'échapper aux contaminations réactives au fait expérimental (B.1.1. et B.2.1.). Il n'est guère possible d'introduire de nouvelles routines, comme l'usage d'une nouvelle grille d'observation et la distribution généralisée d'une pilule, sans que les éducateurs-observateurs et les adolescents-sujets

ne se rendent compte qu'il se passe quelque chose. Dans un établissement semblable à celui où s'est déroulée cette expérimentation, les examens neurologiques et psychologiques sont monnaie courante. Pour ces adolescents, dont certains avaient pratiquement toujours vécu en institution, le test est une routine à laquelle, bien souvent, ils se soumettent sans poser la question du pourquoi, du rythme, etc. Mais, si un beau jour, ils constatent que tous reçoivent une pilule supplémentaire en même temps et aux mêmes heures, on peut s'attendre à une sensibilisation et à une modification des comportements. Quant aux éducateurs, s'étant vu demander de remplir une nouvelle grille, applicable à tous les sujets et exigeant une attention particulière de leur part, ils auront eu leur routine perturbée. En conséquence, ils auront sans doute manifesté et développé de nouveaux types d'interactions avec les adolescents, interactions dont le style aura probablement été limité à la durée de l'expérience.

Bien sûr, tant chez les adolescents que chez les observateurs, ces contaminations ont été effectives pour les deux groupes, expérimental et témoin. Cela ne favorise pas la confirmation des différences entre ces groupes et permet donc de moins facilement soutenir l'hypothèse d'un effet du traitement. Du moins dans ce cas, peut-on penser que le résultat observé est fonction, non seulement du traitement expérimenté (le médicament), mais aussi du traitement secondaire dû à la situation expérimentale (modification des interactions entre éducateurs et adolescents). À la limite, on pourrait se demander si le traitement secondaire n'a pas eu des effets plus puissants que le traitement expérimental. Cela expliquerait l'absence de différence observée : la modification des interactions ayant un effet tel sur les comportements que la modification médicamenteuse supplémentaire ne peut être qu'imperceptible.

Pour échapper à cette contamination, qui par ailleurs conduit dans ce cas à la prudence quant à l'usage du médicament et donc pourrait être considérée comme bénéfique, il eut sans doute fallu appairer les sujets, non pas deux par deux, mais trois par trois, ou plus encore. Mais l'appariement trois par trois eut suffi si nous avions pensé à appairer, sur base des tests de routine, un troisième sujet pris dans une autre section de l'institution où le traitement, le pseudo-traitement et l'observation continue n'auraient pas été effectués, nous limitant, pour ce groupe, au post-test. Dans ce cas, l'information issue des grilles d'observation aurait été perdue et il eut fallu surmonter la difficulté de rendre aveugle le psychologue quant à l'appartenance des sujets à ce troisième groupe. Cependant, si cette difficulté avait pu être surmontée, une observation terminale et des tests psychologiques

en post-test auraient permis de mesurer l'effet du traitement secondaire (les différences entre les résultats du groupe à anticipation, le pseudo-traitement ou placebo, et du groupe sans traitement).

Enfin, en ce qui concerne les autres sources de contamination (B.4. et C.), la méthode des aveugles les a neutralisées dans la mesure où leurs composantes ont été contrôlées. Or nous venons de voir ce qu'il eut fallu faire pour mesurer l'effet des contaminations B.1.1. et B.2.1.

5 Le plan avec groupes à anticipation et double aveugle

Dans les recherches du type envisagé (étude de l'effet d'un traitement, qu'il soit pharmaceutique, psychothérapeutique ou pédagogique), un schéma idéal devrait s'inspirer d'un plan à six groupes dont deux à anticipation, tel que décrit au tableau 9.5, ou à la figure 9.1 lorsque les groupes sont composés par appariement. Mais en outre, ce plan devrait mettre en situation d'aveugle l'observateur qui prend les mesures (initiales, en cours et terminales), les expérimentateurs (éducateurs, enseignants, etc.) qui appliquent les traitements et les pseudo-traitements et les élèves (avec leurs parents) qui devraient ignorer quel est leur statut dans l'expérience.

Un tel schéma n'est sans doute pas encore parfait; il est surtout rarement praticable en éducation, même s'il se rencontre en médecine⁷. L'effectif de la population, à partir de laquelle les différents groupes doivent être extraits, pose un premier problème : plus le nombre de groupes augmente, plus la source doit être grande. Dès lors, dans bien des cas, le chercheur devra sacrifier l'un ou l'autre des groupes. En outre, cela multiplie les observateurs-testeurs de même que les agents qui appliquent le traitement et le pseudo-traitement. Or, plus on recourt à un grand nombre de collaborateurs, plus on risque de vendre la mèche et de transformer les aveugles en semi-voyants. Enfin, cela implique une multiplication des locaux et des agents pédagogiques (administrateurs et enseignants), ce qui rend plus difficile le contrôle des facteurs de l'environnement humain et physique.

Les imperfections qu'un tel schéma peut encore comporter, ainsi que ses difficultés d'application, n'en font qu'un modèle dont on peut s'inspirer. Elles n'en font pas un modèle idéal, car il est trop souvent inaccessible en situation scolaire. D'ailleurs, si le modèle idéal existait, la méthodologie

⁷ Par exemple, Delay J., Perse J. et P. Pichot, *Méthodes psychométriques en clinique*. Paris: Masson, 1966. Le livre II : La validité des tests de personnalité en psychiatrie, p. 93-114, rapporte une très intéressante étude menée selon une variante des plans en double aveugle.

de la recherche cesserait d'être une question de débat, d'enseignement et de préoccupation. Mais en sachant pourquoi et comment de tels plans essaient de limiter les contaminations psychosociales inévitables dans toute recherche en éducation, les chercheurs peuvent anticiper les difficultés de participation des acteurs et d'interprétation des résultats. Enfin, ces difficultés montrent aussi que les plans expérimentaux n'ont qu'une validité interne et externe très limitée et que les chercheurs qui leur accordent encore une préférence, feraient mieux de planifier leur recherche selon des schémas moins ambitieux mais moins trompeurs.

LA MODÉLISATION ET LA SIMULATION

- 1 La modélisation et la simulation : définition
 - 2 Les fonctions et les types de modèles
 - 2.1 *La fonction exploratoire et préparatoire*
Illustration d'une modélisation et d'une simulation préparatoire
 - 2.2 *La fonction interprétative*
 - 2.3 *La fonction didactique*
 - 3 La construction d'un modèle simulable
 - 4 L'application d'un modèle : la simulation
-

1 La modélisation et la simulation : définition

La modélisation et la simulation sont deux phases d'une activité que la plupart des gens pratiquent spontanément sans y mettre ces étiquettes¹. Lorsqu'un adolescent rentre en retard à la maison et qu'il réfléchit à l'excuse qu'il devra présenter à ses parents, il fait des simulations mentales : il évalue l'une après l'autre les excuses possibles en imaginant comment ses parents vont y réagir avant de choisir celle qui a le plus de chances d'être vraisemblable. Pour arriver à faire ce test imaginaire des excuses possibles, cet adolescent utilise une sorte de modèle des réactions de ses parents à ce qu'il peut dire. En se basant sur l'expérience qu'il a de situations semblables, il s'est fait une représentation de leurs réactions qui lui permet d'imaginer les différents scénarios possibles, de prévoir leurs effets, et de choisir ensuite celui qui passera le mieux. Au sens technique des termes, la représentation générale des réactions de ses parents sur laquelle il se base pour imaginer comment ils réagiront à ses excuses constitue le modèle des réactions parentales de cet adolescent, et les différents scénarios qu'il imagine à partir de ce modèle sont des simulations.

Modéliser, c'est construire une représentation générale et simplifiée du réel; c'est ébaucher une caricature, un plan, un schéma, à partir duquel on pourra essayer différentes fictions particulières, ou simulations, de la chose que l'on veut représenter. Au théâtre, le texte écrit par l'auteur propose un modèle des paroles et des comportements de ses personnages. Ce modèle est toujours une sorte de caricature : aucune personne réelle ne se comporte de cette manière. Le portrait suggéré par l'auteur accentue certains traits ou en estompe d'autres afin de mieux mettre en évidence les caractéristiques du personnage qui sont intéressantes pour le déroulement de l'intrigue. Mais le texte de l'auteur, le modèle, reste général : il constitue la base à partir de laquelle différents metteurs en scène n'imagineront pas les mêmes jeux d'acteurs. Plus encore, un même metteur en scène pourra, durant les répétitions, imaginer et essayer plusieurs jeux, plusieurs déplacements des acteurs et du décor, avant de choisir la mise en scène qui sera suivie lors des représentations devant le public. Le metteur en scène part donc d'un modèle, le livret de l'auteur, il fait des simulations, les différentes scénarisations, et il en retient une pour la présentation au public. La simulation constitue donc une concrétisation particulière d'un modèle et,

1 Modélisation et simulation sont les étiquettes techniques ou savantes de ce que d'autres nomment du barbarisme «design» et de ce qui en français se dirait «conception et plan».

comme le modèle est une représentation simplifiée du réel, la simulation est une concrétisation fictive et simplifiée du réel.

Les écrits non spécialisés ne font pas souvent la distinction entre modèle et simulation. Cette distinction est le fait des mathématiciens, des physiciens, des ingénieurs et des économistes qui construisent, à partir de diverses théories, des modèles d'une machine ou de processus². Ces modèles sont constitués d'une série de formules reliant des variables. Plus le phénomène modélisé est complexe, plus le programme (la liste des formules à exécuter les unes à la suite des autres) est long et complexe, plus il requiert un ordinateur puissant. Pour exécuter le programme, il faut donner des valeurs initiales aux variables indépendantes du programme. Le fait d'attribuer des valeurs aux variables et d'exécuter le programme constitue une simulation. Les météorologues utilisent ainsi des modèles mathématiques du climat, et chaque prédiction est le résultat d'une simulation dans laquelle des valeurs particulières (observées à différents endroits du globe) ont été données à des paramètres comme, entre autres, la pression barométrique, les nuages, la vitesse des vents, le pourcentage d'humidité, etc. Quand les ingénieurs préparent le dessin d'une voiture, ils

- 2 Pour une étude approfondie de la modélisation et de la simulation, on peut consulter :
- Popper J., *La dynamique des systèmes. Principes et applications*. Paris : Les Éditions d'Organisations, Eyrolles éditeur ; Montréal : PUQ, 1973.
 - Le Moigne J-L., *Les systèmes de décision dans les organisations*. Paris : PUF, 1974
 - Le Moigne J-L., *La théorie du système général, théorie de la modélisation*. Paris : PUF, 1977.
 - Peaucelle J-L., *Les systèmes d'information, la représentation*. Paris : PUF, 1981.
 - Walliser B., *Systèmes et modèles, introduction critique à l'analyse de systèmes*. Paris : Seuil, 1977.

Ces ouvrages se réfèrent tous à une conception formalisante (mathématique) et mécanique de la modélisation qui ne tient pas toujours compte de certaines caractéristiques fondamentales du comportement humain. Une certaine conception technologique du «design pédagogique», assez mécaniste et behaviorale, est à mi-chemin entre la formalisation des auteurs cités ci-dessus et une conception plus psychosociale. L'ouvrage de Jean Berbaum en est une illustration: il essaye d'appliquer aux actions de formation l'analyse systémique proposée par Jean-Louis Le Moigne.

- Berbaum J., *Étude systémique des actions de formation*. Paris : PUF, 1982.

Enfin, d'autres auteurs ont tenté des applications et des réflexions qui seraient plus pertinentes à la description des systèmes humains. Voir par exemple :

- Checkland P., *Systems Thinking, Systems Practice*. New-York : J. Wiley, 1982.
- Barel Y., *Le paradoxe et le système*. Grenoble : Presses universitaires de Grenoble, 1979.
- Wilden A., *Système et structure. Essais sur la communication et l'échange*. Montréal : Boréal-Express, 1983 (1ère édition anglaise : 1972).

font de même pour les éléments qui constitueront le véhicule. Ils peuvent ainsi essayer sur ordinateur puis sur des maquettes (modèles réduits) les différentes combinaisons de moteur, de carrosserie, de pneus, etc., qui produiront en fin de compte le véhicule attendu par le marché. Si le modèle du véhicule a été bien conçu et si des simulations assez nombreuses et diversifiées ont été faites, les ingénieurs auront moins de chance de se tromper en produisant la voiture que les météorologues en prédisant le temps, car ils peuvent contrôler les différents éléments qui composeront le véhicule.

2 Les fonctions et les types de modèles

La modélisation et la simulation en sciences humaines et en éducation ne peuvent pas atteindre le même degré de formalisation mathématique que celles des physiciens, car ces disciplines ne disposent pas de théories assez élaborées pour construire de tels modèles. Mais on pourrait les utiliser plus souvent. En effet, même à un niveau que certains qualifieraient de sommaire, la modélisation et la simulation ont des propriétés très intéressantes, qui ne sont pas assez exploitées tant dans la recherche que dans le développement et l'action. La modélisation et la simulation peuvent avoir des fonctions exploratoires et préparatoires, interprétatives et didactiques.

Avant d'examiner ces différentes fonctions, il faut aussi noter que la modélisation a un éventail d'applications très large. Les modèles et les simulations peuvent représenter différents objets : des personnes, des groupes, des organisations aussi bien que des outils, des matériaux ou même des comportements et des processus mentaux. Il suffit que l'on puisse imaginer en pensée et remplacer ce qui est à modéliser par des symboles, par des traces perceptibles ou par d'autres objets, pour qu'un modèle puisse être créé. Toute représentation entretenant une similitude (de forme, de structure, de fonction, etc.) avec l'objet à représenter, peut fonctionner comme modèle aboutissant à des simulations. On peut donc trouver différents types de modèles et de simulations, dont par exemple :

- Les modèles mentaux et les simulations imaginaires, non communicables, mais utiles;
- Les modèles et les simulations physiques, comme les maquettes, les modèles réduits;
- Les modèles et les simulations schématiques comme les graphiques, les images, les organigrammes, les listes de procédures;
- Les modèles et les simulations langagiers comme les analogies, les mythes, les fables, les paraboles;

- Les modèles et les simulations théâtrales comme les jeux de rôle, les différentes formes de micro-enseignement;
- Les modèles formels comme les lois, les règles, les algorithmes, les programmes d'ordinateurs, les formules, etc.

2.1 La fonction exploratoire et préparatoire

La modélisation et la simulation d'un objet peuvent avoir une fonction exploratoire parce qu'elles permettent de manipuler la représentation de l'objet comme si c'était l'objet lui-même. Ainsi, on peut voir agir la représentation d'un objet ou voir à distance la représentation d'une action sur cet objet, sans qu'il soit là, sans que cette manipulation n'ait un effet réel. Même sur des maquettes ou des modèles réduits, la modélisation et la simulation restent de l'ordre de la fiction, du simulacre, comme au cinéma ou avec les jeux électroniques. Puisque le concept est aussi une représentation abstraite d'une chose qui permet de l'utiliser dans des énoncés, les modèles conceptuels dérivés des théories sont des représentations qui permettent de concevoir l'action des choses ou des individus ou l'effet d'une action sur des événements, sans que les choses et les individus aient été manipulés réellement et directement. La modélisation permet ainsi de projeter, dans des simulations, des constructions ou des actions, de planifier les interventions et d'organiser la mise en place des moyens nécessaires à leur réalisation. La simulation permet d'en prévoir les effets potentiels, c'est-à-dire de les évaluer dans des contextes précis bien que fictifs, c'est-à-dire sans avoir à effectuer réellement ces tâches projetées et sans devoir réellement construire les différents contextes possibles. La modélisation et la simulation constituent donc des outils très puissants d'exploration, de préparation, de choix et de décision avant l'action. Les modèles réduits et les maquettes d'avions et de carrosseries d'automobiles que les constructeurs étudient en soufflerie illustrent cette fonction exploratoire et préparatoire. La simulation permet aussi d'évaluer des projets qui ne seront jamais effectués, parce que ses résultats conduisent à rejeter ces projets pour des raisons morales, politiques ou économiques. La simulation permet en effet de montrer l'effet désastreux de certains projets avant et parfois sans que le désastre n'ait à se produire. Ainsi, dans certaines études environnementales, les effets des marées et des courants sur la propagation de pollutions éventuelles ont pu être étudiés sur des maquettes installées dans des bassins représentant des régions côtières. Ces simulations ont permis d'anticiper les moyens pour protéger ces régions au cas où un accident arriverait.

Bien des erreurs seraient évitées si l'on disposait du temps nécessaire pour faire des simulations exploratoires. Par exemple en recherche, la modélisation et la simulation permettent de mieux planifier les différentes phases de la recherche, de prévoir quand on devra aller sur le terrain, de corriger le calendrier pour ne pas y arriver pendant les vacances scolaires, de savoir quand il faudra engager certains auxiliaires ou assistants et de réserver à temps les spécialistes qu'il faudra consulter. La modélisation et la simulation des projets peuvent être facilitées par des logiciels³ qui représentent graphiquement les étapes et les phases d'un projet et qui calculent les effets de la modification d'une étape ou d'une partie du plan ainsi que les effets d'un changement dans l'allocation des ressources.

Illustration d'une modélisation et d'une simulation préparatoire

Partant de la figure 8.1 qui constitue un modèle schématique d'une recherche exploratoire (la liste de ses phases), la séquence des principales phases s'exprimerait comme dans le tableau 11.1.

- | | |
|-----|--|
| 1 | Description du problème <i>in situ</i> |
| 2 | Explicitation des préconceptions du chercheur et des acteurs |
| 3 | Élaboration du cadre conceptuel et méthodologique |
| 3.b | Préparation des outils de constitution des données |
| 4 | Constitution des données |
| 4.b | Analyse de la valeur des données |
| 5 | Analyse descriptive des données |
| 6 | Synthèse et traitement des données |
| 7 | Interprétation des résultats et formulation d'hypothèses |
| 8 | Validation des résultats et des hypothèses |
| 9 | Rédaction du rapport |

Tableau 11.1 : Modèle de la séquence des principales phases d'une recherche exploratoire.

Chacune de ces phases pourrait être décomposée et faire l'objet d'un modèle partiel. Ainsi la description du problème, soit la première phase, peut se décomposer en six éléments comme dans le tableau 11.2. Mais les énoncés de ce tableau ne précisent que les résultats de cette phase, c'est-à-

3 Par exemple, le logiciel MacProject, édité par Claris, permet des simulations dans la gestion de projets.

dire qu'on y trouve les six éléments (ou étapes) auxquels doivent aboutir les phases ou procédés qui permettent de produire la description du problème.

- | | |
|-----|--|
| 1.1 | Énoncé local dans les catégories et les perspectives des acteurs |
| 1.2 | Contexte de la situation dans laquelle baigne le problème |
| 1.3 | Implications du problème sur les acteurs et le milieu |
| 1.4 | Conséquences du problème |
| 1.5 | Priorités à envisager lors de la résolution du problème |
| 1.6 | Contraintes du contexte délimitant les solutions du problème. |

Tableau 11.2 : Liste des étapes d'une description du problème *in situ*.

Chacune de ces étapes doit être décomposée en termes de phases qui y conduisent, c'est-à-dire qu'il faut préciser qui (agent), avec quels moyens (ressources), fera quoi (processus ou procédés), dans quel délai (début et fin des opérations), pour produire quoi (résultat). Cette décomposition des étapes peut déboucher sur une simulation lorsque le chercheur, ayant choisi son problème (par exemple, le décrochage scolaire) fait les estimations suivantes.

A – L'élément 1.1 exige qu'un enquêteur se rende sur le terrain et effectue des entrevues non structurées, de type conversation, afin d'écouter comment le problème se présente selon les différents acteurs.

A1 – Préalablement à ces entrevues, il faut

A2 = identifier un terrain susceptible d'être ouvert (une école où le problème existe, dont des membres du personnel ont accordé des entrevues aux journaux locaux, etc.);

A3 = identifier les informateurs (quels sont les acteurs impliqués, à quel degré);

A4 = obtenir les autorisations des enseignants, de la direction, des adolescents et de leurs parents;

A5* maîtriser les techniques de l'entrevue non structurée et de l'enregistrement de ce type de matériel, sinon ou bien s'entraîner auprès d'un expert, ou bien engager un expert;

- A6 – prévoir et organiser la présence sur le terrain afin de rencontrer tous les informateurs nécessaires;
- A7 – à la suite de ces entrevues, il faudra effectuer un codage ouvert, dans une perspective phénoménologique et sociale puisque le phénomène (le décrochage) prend son sens dans la perception des relations (adolescents – école – enseignants – parents);
 - A8* connaître les principes de l'analyse phénoménologique, sinon engager un expert en phénoménologie;
 - A9* connaître les principes, par exemple, de l'interactionnisme symbolique et ses méthodes de travail, sinon engager un expert.

B – L'élément 1.2 exige une enquête à propos des conditions de vie et de l'organisation des activités dans l'école, à propos du contexte social, économique et culturel des décrocheurs et des non-décrocheurs de l'école. Cette enquête implique

- B1 – une observation participante sur le terrain pour en saisir la culture;
 - B2* maîtriser cette technique ou d'engager un expert;
- B3 – une recherche de données socio-économiques auprès d'organismes régionaux qui disposent de l'information;
- B4 = identifier les organismes;
- B5 = préparer la liste des questions;
- B6 = prendre les rendez-vous;
- B7 = structurer le matériel recueilli;
 - B8* maîtriser les techniques d'analyse de contenu à partir de listes de rubriques, sinon d'engager un expert;
- B9 – des entrevues structurées ou des questionnaires auprès de chaque adolescent et de ses parents;
- B10 = construire des grilles d'entrevue et des questionnaires puis les valider.

C – Les éléments 1.3 à 1.6 peuvent s’obtenir à partir des entrevues et des questionnaires à condition de :

C1 = prévoir des questions qui s’y rapportent et des rubriques qui les couvrent dans la grille d’analyse;

C2* maîtriser les techniques d’analyse de contenu à partir de listes de rubriques, sinon d’engager un expert (compétence identique à celle de la ligne B8).

Mais la simulation n’est alors qu’ébauchée : il faudrait encore y apporter des précisions avant de pouvoir envisager les autres possibilités reliées à la planification du projet. Ainsi, la simulation débouchera sur des planifications différentes selon que le chercheur est un spécialiste de la recherche, qu’il est subventionné ou non, et qu’il maîtrise ou non la plupart des compétences (marquées d’un *) indispensables à la réalisation du projet, ou selon qu’il s’agit d’un étudiant qui ne maîtrise pas ces compétences et qui ne jouit pas de subventions lui permettant d’engager les experts qui suppléeraient à son ignorance. S’il s’agit d’un étudiant commençant des études de deuxième cycle, il devra ajouter des phases préalables (des cours ou autres activités d’apprentissage) à inscrire dans son programme de scolarité avant de passer aux phases d’élaboration des outils de recherche et d’entreprendre les démarches pour accéder au terrain. Une fois que l’analyse des phases (déroulement d’activités) et des étapes (résultats intermédiaires) du projet est complétée, en général en débutant l’analyse par l’étape terminale et en lui fixant un délai final, il suffit de reconstruire le cheminement en partant des phases préalables, pour aboutir à une simulation du projet comme celle illustrée à la figure 11.1.

Cette figure 11.1 suggère que la simulation, ébauchée par l’étudiant, ne pourra être efficace que s’il dispose déjà de modèles suffisamment précis pour envisager les phases indispensables à la réalisation de la recherche. L’importance du rôle de tuteur ou de directeur de recherche apparaît évidente.

En arrêtant à ce niveau l’illustration de la fonction de préparation que peut jouer une simulation, on aperçoit déjà les embûches qu’elle indique et les points critiques auxquels il faudra être attentif dans la planification et la réalisation du projet. Une simulation exploratoire et préparatoire bien faite indique toutes les exigences d’un projet, entre autres, les compétences à maîtriser, les ressources dont il faut disposer ainsi que les moments où elles doivent être disponibles afin que le projet puisse se terminer en temps

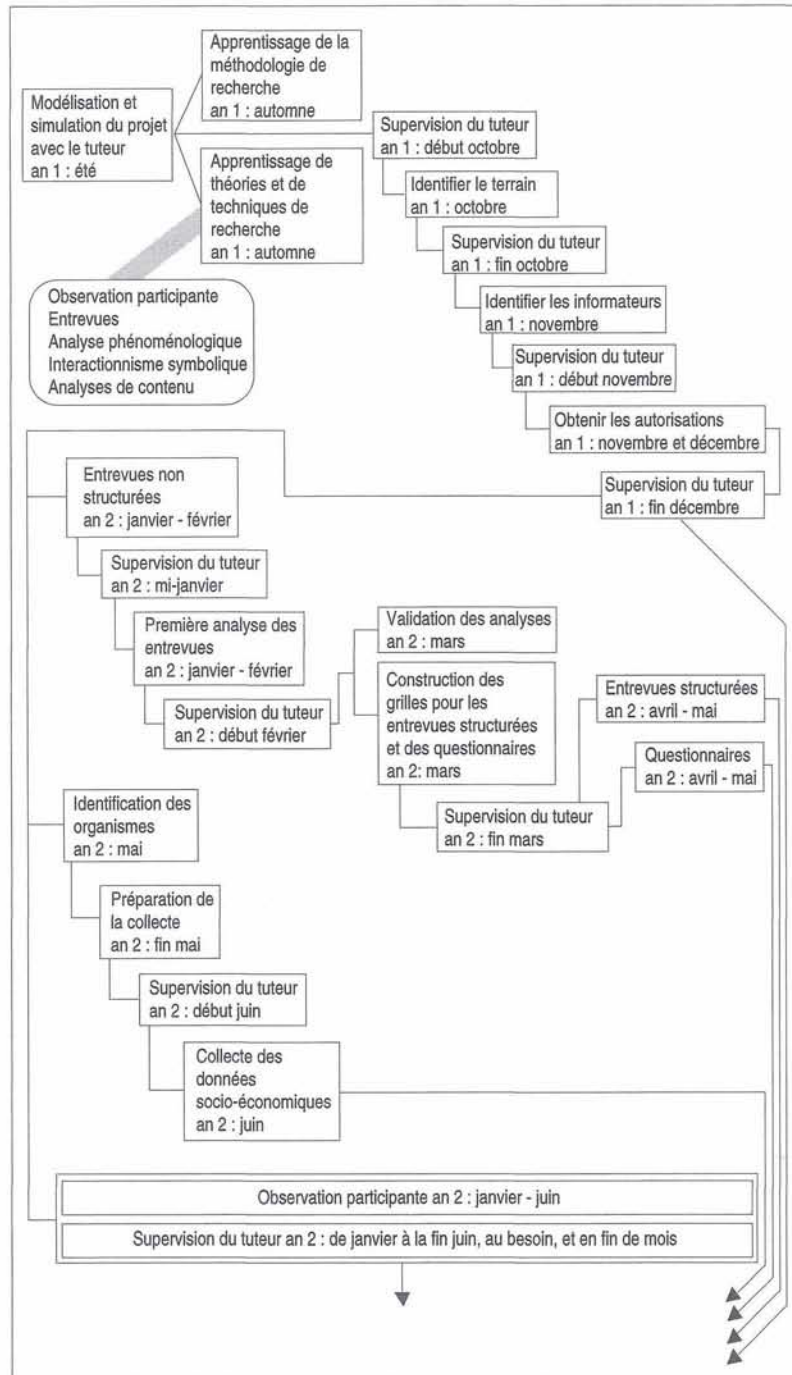


Figure 11.1 : Organigramme partiel d'une simulation des phases d'un projet.

voulu. La simulation exploratoire permet donc de préparer des demandes et d'évaluer de façon réaliste dans quelle mesure un projet est faisable. Cela peut avoir pour effet de décourager avant d'entreprendre mais, si tel est le cas, cela évite aussi de s'engager tête baissée dans un cul-de-sac. Si, par contre, les points critiques paraissent surmontables, la réalisation de la recherche sera mieux planifiée et se déroulera plus aisément.

Il faut encore préciser que toute simulation reste vivante : un organigramme comme celui de la figure 11.1 peut toujours être modifié parce qu'il n'est qu'une simulation. Il semble que plusieurs refusent de faire ce travail de simulation, assez fastidieux, craignant d'y être liés alors qu'ils ont l'intuition, très juste il est vrai, qu'un projet de recherche exploratoire évolue et ne peut être complètement planifié dès le départ. Mais si la préparation du projet se fait dans un esprit de simulation, le chercheur garde la liberté de s'adapter à ce qu'il découvrira en cours de route et de modifier la suite de la simulation.

2.2 La fonction interprétative

Les modèles ont aussi une fonction interprétative dans la mesure où l'on peut emprunter une représentation d'un phénomène connu et accessible pour rendre compte de phénomènes peu connus ou inaccessibles. Ainsi, Rutherford-Bohr s'est inspiré de la structure du système solaire (le soleil entouré des planètes) pour interpréter les phénomènes atomiques (le noyau au centre et les électrons autour). De même, Freud s'inspire du mythe d'Oedipe pour proposer une explication des rapports entre la construction de la personnalité et les relations interpersonnelles. Cette fonction interprétative du modèle repose sur l'analogie et suppose un certain isomorphisme, c'est-à-dire une ressemblance des formes entre les phénomènes A et B. Mais cet isomorphisme (égalité des formes ou, par extension, des fonctions) ne doit pas être parfait pour être fonctionnel. Lorsque les chercheurs tentent d'analyser leurs données ou d'interpréter leurs résultats, la démarche la plus féconde consiste la plupart du temps à chercher à quelle forme connue ressemble le matériel obtenu.

La force des méthodes quantitatives réside d'ailleurs dans le fait que les mathématiques ont proposé différents modèles de distribution (de formes) des données ou des résultats pour lesquels des paramètres ou, autrement dit des caractéristiques mesurables, sont connus. Dès lors, une interprétation se fait plus facilement, sinon automatiquement : il suffit de repérer à quelle forme de distribution mathématique (par exemple, la courbe normale, une

loi de Poisson, chi-carré, etc.) ressemble le matériel obtenu pour lui attribuer les caractéristiques de la distribution mathématique. La lacune de telles interprétations tient à deux raisons. D'une part, l'écart est souvent excessif entre le modèle formel proposé par les mathématiques et les caractéristiques fondamentales de l'objet représenté. D'autre part, la similitude des formes suggère une similitude des fonctions, mais cette relation entre les similitudes s'avère trop souvent illusoire. En effet, bien des choses peuvent avoir une apparence semblable et agir, c'est-à-dire fonctionner, différemment. Ce que la sagesse populaire exprime en disant que les apparences sont trompeuses. Les chercheurs doivent accepter que la valeur interprétative des modèles, qu'ils soient quantitatifs ou qualitatifs, soit limitée à leur fournir une première piste analogique à partir de laquelle ils peuvent réfléchir et chercher les interprétations qui seront pertinentes à l'objet, au problème étudié et au contexte dans lequel il s'est posé.

2.3 La fonction didactique

Les modèles peuvent aussi avoir une fonction didactique. En effet, une manière de faire percevoir en un coup d'oeil des situations (ou des processus) complexes ou des phénomènes qui se déroulent sur des périodes assez longues de temps, consiste à en réaliser un modèle conceptuel d'un niveau assez élevé de généralité, que l'on peut présenter sous la forme d'un plan, d'un schéma ou d'un organigramme. Les différents schémas qui ont été proposés dans les chapitres précédents (par exemple, pour la recherche développement, la dissémination de l'innovation, l'évaluation L.V.R. ou pour la recherche nomothétique) sont des modèles qui permettent de se rendre compte en un instant d'une série d'opérations ou de phases qui se déroulent dans le temps. De plus, la comparaison de ces modèles permet aussi de comprendre ce que ces types de recherche ont en commun et ce qui les différencie. Une fois que la vue d'ensemble du modèle a été comprise, on peut ensuite proposer des modèles plus détaillés, d'un niveau plus analytique, permettant un approfondissement de chacune des phases ou de chacun des éléments du modèle général. La simulation didactique peut alors, comme on le fait dans les illustrations et les schémas incorporés aux manuels, ne retenir que les éléments sur lesquels on veut attirer l'attention, en estompant ou en éliminant de la représentation les détails qui pourraient constituer des distractions.

Mais la fonction didactique des modèles est plus large encore : tout l'enseignement repose sur la construction de modèles de ce que sont les situations et les actions de la vie quotidienne, et sur leur simulation. En

effet, un des avantages de la vie scolaire est de confronter les élèves à des simulations de ce qu'est la réalité. Cela permet un apprentissage dans lequel les erreurs peuvent ne pas avoir de conséquences irrémédiables comme dans la vie. Cela permet aussi de découper la complexité de gestes afin de les apprendre morceau par morceau, puis de les intégrer une fois que chaque partie est assimilée. Sans que cela soit nommé ainsi, ni fait de manière systématique, bien des apprentissages professionnels ne pourraient pas se réaliser sans risques pour la santé des étudiants s'ils n'avaient pas été d'abord modélisés et simulés. Les simulateurs de vol utilisés pour l'entraînement des pilotes en sont une illustration, le micro-enseignement en est une autre.

3 La construction d'un modèle simulable

Pour être efficaces dans la recherche, les modèles doivent avoir été composés à la suite d'une analyse minimale de l'objet. Selon les cas, les modèles sont déduits ou induits. Déduits, ils sont construits à partir d'une conception ou d'une théorie provisoire applicable à l'objet; induits, ils proviennent d'observations et d'enquêtes fournissant une description préliminaire de l'objet.

Cette description initiale du modèle suppose que l'objet soit analysé afin d'en identifier les éléments principaux, c'est-à-dire ceux qui sont opérants dans le problème concerné. On parle d'éléments opérants (ou de variables) lorsque l'on effectue la distinction entre, d'une part, les éléments de l'objet que l'on peut identifier, contrôler ou manipuler et auquel on peut attribuer des valeurs (quantitatives ou qualitatives) et, d'autre part, les éléments non manipulables, non contrôlables et non mesurables, éléments qui restent le plus souvent hypothétiques. Seuls les éléments identifiables, mesurables et contrôlables ou manipulables sont considérés comme opérants et doivent, dans la mesure du possible, être insérés dans le modèle de l'objet. Les éléments hypothétiques sont, en principe, évités. Cependant, ils peuvent être intégrés au modèle s'ils proposent un principe ou une règle hypothétique permettant de comprendre ou d'expliquer les liens entre des états ou des éléments opérants qui ne pourrait pas s'expliquer autrement.

Pour illustrer cette distinction, on peut imaginer un modèle pédagogique des relations entre le professeur et ses élèves. Si, à l'occasion d'une recherche, nous avons à agir sur le système ainsi modélisé, il apparaît assez vite que certains de ses éléments, comme les parents des élèves et leur milieu culturel, ne sont pas des éléments manipulables; mais ils sont observables et mesurables (au moins qualitativement). Ils constituent des

éléments opérants de ce modèle, d'autant plus opérants qu'il est impérieux d'en tenir compte. Par contre, la pulsion ou le désir de savoir (au sens psychanalytique) sont des concepts purement hypothétiques, probablement intéressants (aux yeux des psychanalystes) pour expliquer une notion comme la motivation des élèves au travail scolaire. Mais ces concepts psychanalytiques ne sont guère utiles dans le modèle, car ils ne sont ni observables ni manipulables par le chercheur (ni par l'enseignant). Ils ne seront pas intégrés au modèle et ils y seront remplacés par d'autres concepts comme la motivation ou les attentes dans la mesure où l'on dispose d'instruments permettant de les observer ou de les mesurer et de théories qui suggèrent des actions pour les manipuler.

À partir de la théorie du système général, présentée par Le Moigne⁴ comme une théorie de la modélisation, les éléments opérants peuvent être envisagés sous trois dimensions :

- Leur état, c'est-à-dire leur valeur, leur grandeur ou leur rôle, leur fonction;
- Leur structure, c'est-à-dire leur position et leurs connexions ou leurs relations avec les autres éléments du modèle ou avec les sous-éléments qui les composent. La structure est envisagée tant sous ses aspects internes qu'externes;
- Leur évolution, c'est-à-dire leurs activités sur eux-mêmes et sur d'autres éléments, ce que l'on peut observer en termes de productions, de transformations, de transports et de contrôle (décisions et régulations réciproques ou non).

Idéalement, l'élaboration d'un modèle, qui soit utilisable pour des simulations, examine et délimite quelles sont les variations possibles de chacune de ces dimensions, quels facteurs pourraient provoquer une variation, et quelles seraient les règles de ces variations.

- De manière plus précise, il faut fixer les limites de variations en identifiant l'éventail des états (minimum et maximum possibles des valeurs, variétés des fonctions et des rôles), en décrivant les formes admises des structures et en définissant les types d'évolutions permises. Dans certains cas, les variations peuvent être réduites à l'existence ou non de l'élément (variation de zéro à un); dans d'autres cas les variations seront très étendues. Parfois les variations sont très bien décrites par des théories, parfois on en ignore tout.

4 Le Moigne J.L., *La théorie du système général, théorie de la modélisation*. Paris : PUF, 1977.

- Ensuite, il faut identifier les facteurs de variations, c'est-à-dire les relations de concomitance, d'antériorité ou de conséquence entre les variations d'un élément et les variations d'un ou de plusieurs autres éléments.
- Enfin, une fois les limites et les facteurs de variations établis, on précise les règles de variations de chacun des éléments en ce qui concerne les paramètres de forme, de sens, d'intensité ou d'amplitude et de fréquence des variations. La précision formelle des règles dépend évidemment des connaissances théoriques du domaine.

Variations / dimensions	Etat	Structure	Evolution
Limites de variation	?	?	?
Facteurs de variation	?	?	?
Règles de variation	?	?	?

Tableau 11.3 : Matrice de simulation pour un élément d'un modèle.

Ainsi, une fois identifié chaque élément opérant du modèle de l'objet, il faudrait remplir les cellules de la matrice des dimensions et des variations, ne serait-ce qu'en maintenant ouverte une interrogation. La matrice du tableau 11.3 illustre à quel travail l'analyse devrait aboutir afin que le modèle puisse être utilisé comme base pour des simulations, chaque simulation exprimant le résultat de l'attribution d'une valeur précise dans chacune des cellules de la matrice. Si, par exemple, on veut construire un modèle pour des simulations du développement d'un matériel didactique, le premier élément à modéliser serait le client utilisateur, ou le public cible selon la figure 7.5. Prenant le cas du développement d'un matériel utilisant une ou plusieurs démarches pédagogiques afin d'apprendre à des étudiants des stratégies de lecture efficaces, la matrice de simulation du client pourrait se discuter comme suit, chacune des cellules pouvant être déterminée à l'intérieur d'antipodes.

Par exemple, l'état pourrait varier du jeune étudiant de premier cycle à l'étudiant adulte qui revient aux études après vingt ans d'enseignement en première année du primaire, cet étudiant pourrait être très cultivé ou non, il pourrait vouloir apprendre des stratégies de lecture ou vouloir les améliorer. Sur le plan structurel, le client pourrait travailler seul ou en équipe (structure externe) et disposer de connaissances plus ou moins élaborées sur l'analyse des discours, la rhétorique, la logique, la syntaxe (structure interne). En termes d'évolution, les stratégies de lecture pourraient n'être

acquises qu'au niveau de l'information, ou de l'application et du transfert sur tout type de discours; elles pourraient aussi rester une tâche volontaire, devenir des automatismes ou atteindre le niveau d'habiletés métacognitives. C'est à l'intérieur de ces limites de variations qu'un développeur éventuel devra faire ses choix, bien que ces variations sont peut-être reliées à des facteurs déterminants. Dans cet exemple, l'état pourrait dépendre de l'âge des étudiants visés, la structure interne de la filière scolaire déjà suivie par l'étudiant et l'évolution du type de pratique professionnelle à laquelle l'étudiant se destine. Les règles de variations laisseront sans doute vides (?) quelques cellules : en ce qui concerne l'état, peu de règles sont disponibles alors que, pour la structure externe, une correspondance terme à terme pourrait être retenue entre travaux en équipe et matériel graphique ou audiovisuel d'une part, entre travaux isolés et matériel informatique d'autre part.

Chaque élément du modèle du développement de ce matériel (le contenu, la stratégie d'enseignement, la forme de présentation, etc.) devrait être traduit sous forme d'une matrice de simulation avant que des simulations efficaces ne soient entreprises. C'est souvent l'absence d'une telle analyse qui conduit les concepteurs à produire des matériels didactiques trop linéaires, peu vivants et qui semblent manquer d'imagination : sans matrice de simulation, ils n'ont pas vu l'éventail des possibles et des combinaisons entre ces possibles, ils en sont restés à un design monolythique.

Le modèle ainsi obtenu ne doit pas nécessairement être systémique : tous les modèles ne sont pas composés d'éléments en interaction réciproque. Dans les modèles les plus pertinents à certains objets, les éléments sont en séries parallèles ou en arborescences (embranchements qui se multiplient comme les branches d'un arbre). Par exemple, l'arborescence est typique du modèle de la communication et du pouvoir dans les organisations militaires ou dans l'Église catholique. Le modèle de la recherche exploratoire, esquissé au tableau 11.1, est un schéma en forme de liste de phases qui, lorsqu'on le détaille, devient une liste plus allongée de productions et de décisions impliquant certaines actions, comme on en a simulé une partie à la figure 11.1. Les interactions et les rétroactions sont rares dans ce genre de modèle qui représente des procédés et qui se déroule de manière plus linéaire qu'en boucle. Il n'y a pas de prototypes généralisables de modèles, même si le modèle systémique semble avoir eu un certain succès et une relative efficacité lorsqu'il s'agit de représenter des organismes complexes comportant de nombreux sous-ensembles aux fonctions diverses.

4 L'application d'un modèle : la simulation

Le modèle d'un objet a une valeur générale pour cet objet; il en constitue une représentation non contextualisée, hors du temps et de l'espace. Le modèle, c'est la charpente qui permet de reconnaître un objet sous diverses apparences possibles et qui permet d'imaginer l'allure particulière qu'il prend dans des contextes spatio-temporels précis (lorsque l'on fixe des attributs définis à chacun des éléments manipulables). C'est en référence à un modèle qu'on peut comparer et évaluer chacune des apparences possibles de l'objet, chacune de celles-ci étant aussi une apparence particulière, ou une concrétisation, du modèle. Chacune des apparences du modèle obtenue en fixant une valeur à un ou plusieurs éléments que l'on peut manipuler, constitue une simulation. Chaque simulation est une fiction contextualisée de l'objet, et ce sera la comparaison de différentes simulations d'un objet qui permettra de choisir quelle en sera la meilleure forme et quelle sera la stratégie la plus efficace pour y arriver. Mais la comparaison et l'évaluation des différentes simulations exigent que soit respecté le caractère général du modèle, c'est-à-dire que l'on maintienne la stabilité de la charpente. Aussi longtemps que des simulations sont en cours d'essai, on ne peut pas changer la matrice de simulation, c'est-à-dire les limites des valeurs attribuables, les facteurs de variations et les règles de variations tant pour les états, pour les structures, que pour les évolutions envisagées dans la description initiale du modèle. Si l'on modifie la matrice de simulation, on introduit des exceptions dans les comparaisons, ce qui empêche d'évaluer correctement.

Implicitement, la plupart des individus font de la simulation élémentaire sans le savoir. Ainsi, lorsqu'un enseignant évalue mentalement l'effet de certaines interventions qu'il projette de faire à l'égard d'un de ses élèves difficiles, il a en tête un modèle de l'interaction entre cet élève et lui-même. Ce modèle de l'interaction comporte plusieurs sous-modèles, tels ceux de l'élève et de ses conduites, des interventions possibles, des significations que l'élève attribuera aux diverses interventions, des interprétations des réactions de l'élève que l'enseignant peut faire, des réactions qu'il peut tolérer, etc. Ce modèle, plus ou moins raffiné, lui permet de construire les différents scénarios correspondant aux interventions possibles, c'est-à-dire de simuler les interventions projetées, leurs effets sur l'élève, les réactions de ce dernier et sa propre capacité de répondre aux gestes de l'élève.

Mais de telles simulations mentales, si elles sont très utiles pour leur auteur, ne sont pas communicables. Elles ne permettent pas d'échanger avec d'autres les évaluations que l'on a pu faire, elles ne permettent pas de garder une trace des calculs mentaux que l'on a faits et donc d'apprendre de ces expériences. Le recours explicite à la modélisation et aux simulations que l'on traduit de manière communicable (schémas avec règles permettant d'établir les valeurs à observer) permet de discuter avec des collègues, avec des chercheurs ou avec des utilisateurs potentiels, et donc d'ajuster la préparation de l'action à l'ensemble des acteurs concernés. De plus, cela permet aussi de partager l'apprentissage que l'on effectue déjà lors de la comparaison des différentes simulations.

La difficulté de réaliser des simulations tient beaucoup à leur coût. Une simulation implique une étude minutieuse de l'objet pour le modéliser et pour tenter des simulations suffisamment variées afin qu'elles permettent vraiment de comprendre ce qui se passe et de choisir. Or, de telles études consomment souvent beaucoup de temps et exigent un minimum de compétences. Cependant, lorsque l'on s'engage dans une recherche, appliquée ou fondamentale, la préparation du cadre conceptuel ou du cadre théorique implique nécessairement des analyses qui pourraient facilement, et sans un surplus important de travail, déboucher sur une modélisation et sur des simulations. Étant donné les avantages à long terme que la simulation procure en permettant d'éviter bien des erreurs coûteuses dans la réalité, on devrait plus souvent exploiter cette stratégie avant de passer à l'action.

ÉCRIRE LA RECHERCHE

- 1 Une tâche inéluctable
- 2 La nécessité épistémologique
- 3 Le caractère culturel et donc collectif de la recherche
- 4 Les nécessités stratégiques de l'écriture
 - 4.1 *Écrire à d'autres avec leur style*
 - 4.2 *Écrire à propos des énoncés des autres*
 - 4.3 *Utiliser les écrits des autres*
 - 4.4 *Ne pas détourner les alliances ni se mettre hors de portée*
 - 4.5 *Distiller le doute*
 - 4.6 *S'inscrire dans une culture professionnelle*
 - 4.7 *Proposer, non pas des solutions, mais des outils pour leurs solutions*
- 5 Les questions techniques
 - 5.1 *La manière d'écrire : reformuler la pensée pour la communiquer*
 - 5.2 *Savoir quel est le destinataire et s'adresser à lui*
 - 5.3 *Les phases de la recherche et le style d'écriture*
 - 5.3.1 Le problème
 - 5.3.2 Le cadre conceptuel ou théorique
 - 5.3.3 La chronique
- 6 Conclusion

1 Une tâche inéluctable¹

Faire de la recherche, c'est essentiellement écrire. L'écriture prend place tout au long de la recherche et devra répondre à différentes exigences. Nous examinerons d'abord les nécessités épistémologiques et stratégiques de l'écriture, avant d'en aborder quelques aspects plus techniques, liés aux phases de la recherche.

Au préalable, il faut préciser que la recherche dont il est question ici est celle que nous appelons la «recherche ordinaire». Par recherche ordinaire, nous entendons la recherche effectuée par des étudiants de deuxième et troisième cycles, par des chercheurs professionnels ou par des professeurs-chercheurs, dans le cadre de travaux dont les bornes sont fixées par des institutions : règlements pédagogiques, normes institutionnelles, organismes subventionneurs et commanditaires, etc. Nous ne parlerons donc pas de l'écriture des «oeuvres» qui habituellement prend place dans une seconde partie de la carrière, lorsque les chercheurs ont assis leur réputation et se permettent d'écrire en dehors des règles institutionnelles.

La recherche ordinaire n'est probablement pas celle qui révolutionnera, ne serait-ce que quelques dizaines d'années, les manières de voir et de penser de la population. Ce sont les grands discours mythiques qui parviennent à le faire. Le vingtième siècle a ainsi connu le marxisme, le freudisme et le structuralisme, mais comme tous les grands discours mythiques, ils ne tiennent qu'un temps, celui pendant lequel les images qu'ils contiennent correspondent à des valeurs problématiques dans les sociétés où ils ont eu audience.

La recherche ordinaire est celle qui produit les éléments, parfois de micro-éléments qui, enchâssés et confrontés les uns aux autres, finissent par constituer la base de la connaissance à partir de laquelle les gens peuvent réfléchir et agir. En son principe, la recherche vise une chose : mettre en cause le sens commun, la doctrine, le bon sens, les apparences, de façon à libérer la pensée du terrorisme des raisons toutes faites, irréfléchies et imposées d'autorité par le passé et par les détenteurs du pouvoir et des fausses sagesses. Quant à la recherche appliquée, la production de nouveaux outils, d'instruments, d'actions relève de la même préoccupation : améliorer (et donc avoir d'abord critiqué) les instruments mis à la disposition d'une

1 La partie centrale de ce texte (les nécessités épistémologiques et stratégiques) provient d'une conférence prononcée le 15 mai 1990 à l'occasion du colloque «Écrire la recherche» organisé par l'Association pour la Recherche Qualitative (A.R.Q.) dans le cadre du 58e congrès de l'ACFAS à l'Université Laval, Québec.

société. La recherche est donc une reconstruction par interrogation, sinon par contestation. Chaque recherche particulière n'est qu'une brique qui s'ajoute aux autres dans un travail collectif où certains éléments seront conservés et renforcés, d'autres seront restaurés ou modifiés alors que quelques-uns paraîtront nouveaux.

2 La nécessité épistémologique

S'il y a une nécessité épistémologique d'écrire la recherche ordinaire, elle tient à une seule exigence : il n'y a pas de recherche scientifique s'il n'y a pas de trace de la recherche. L'exigence d'écrire la recherche scientifique découle des différences de la science ordinaire avec la conscience de l'expérience, la doctrine révélée et les coutumes, les routines ou les normes sociales.

La recherche scientifique se différencie de la conscience de l'expérience qui n'a pas besoin d'être communiquée. L'expérience, ou du moins la conscience de l'expérience, peut constituer un savoir personnel, idiosyncratique. Si la conscience de l'expérience évolue et peut transformer la conscience elle-même, elle n'entre pas dans un jeu de controverse, car elle cherche une certaine unité, une intégration des expériences, y compris l'atténuation, sinon l'élimination, des dissonances.

La recherche scientifique se différencie aussi de la doctrine révélée des religions ou de certaines philosophies, dans la mesure où ces dernières sont en quête d'absolu et expriment une transcendance immuable dont seules les formes d'expression évoluent avec le temps et avec les cultures. Mais l'intégrisme dogmatique a les mêmes fonctions et condamne toutes les contestations et toutes les déviations, quels que soient le nom et l'origine culturelle des doctrines. La coutume, la routine et la norme sociale, le bon sens et le sens commun se différencient aussi de la recherche scientifique par la lenteur de leur évolution et par le fait que leur remise en question conduit à la marginalisation. Comme la révélation, la norme est une réponse, elle n'est pas une question. L'une et l'autre peuvent s'écrire pour mieux être diffusées et disséminées ou pour mieux contraindre grâce à une codification. Mais elles n'ont pas besoin d'être écrites, il leur suffit de prêtres et de juges.

La recherche productrice de la science ordinaire pose des questions autant à propos de l'ordre que du désordre des événements et de leur apparence, autant à propos de leurs régularités que de leurs ruptures. Dans la mesure où la recherche scientifique ne se laisse pas prendre au piège

d'une illusoire vérité, toute question résolue entraîne une autre question et réinterroge l'objet. Dès lors, l'élaboration des connaissances et des savoirs ne produit que des énoncés provisoires, hypothétiques, conditionnels aux contextes de leur production. Pour que la science ordinaire puisse évoluer de question en question, les recherches doivent pouvoir être analysées et contestées, c'est-à-dire communiquées sous un format qui permette ces analyses. Si la science n'évolue plus, si la recherche ne rebondit pas sur d'autres questions, le savoir se cristallise et n'est plus que de la science toute faite à l'usage des manuels : alors il n'est plus question de recherche.

L'enjeu de l'écriture dans la recherche dépasse de loin la reproduction scolaire, l'utilisation sociale et l'application technologique des résultats. Il tient à un trait fondamental de la recherche scientifique consistant à ne produire que des énoncés contestables, objets de controverses et évolutifs. La recherche ordinaire produit des histoires vraisemblables qui restent des histoires et qui ne font pas ni ne sont l'Histoire.

Or, les questions de la recherche en sciences humaines portent sur des événements inscrits dans le temps et dans l'espace. L'événement, objet de recherche, n'est intéressant que dans la mesure où il est dynamique, où il s'y passe quelque chose, où des échanges se produisent. Mais cette historicité de l'événement le rend fugace, et il disparaît dans le souvenir, à moins qu'on en garde une trace.

La recherche implique donc un double niveau d'écriture. Non seulement les énoncés à propos des événements doivent être écrits, mais l'événement lui-même ne peut pas faire l'objet d'énoncés s'il n'a pas été écrit. Sans une trace, l'objet, le sujet ou l'événement restent éphémères et on ne peut rien en dire qui soit crédible. Sans une empreinte de l'événement, on ne peut que l'avoir vécu, on peut seulement en construire une représentation rétrospective. Dès lors, il n'est que souvenir. Or, on sait que le souvenir enjolive ou dramatise lorsqu'il ne nie pas les événements du passé. En conséquence, pour que l'événement puisse faire l'objet d'énoncés, il faut mettre en place un système d'inscription qui permette de conserver une représentation du déroulement de l'événement tel qu'il s'est passé. Cette trace ou inscription, une fois obtenue, prend la place de l'événement éphémère : la trace reste, alors que l'événement est passé. C'est à partir de cette trace et à propos de cette trace que des énoncés pourront être formulés d'abord comme légende, comme commentaire ou comme description de la trace, ensuite comme interprétation de la description. L'objet des énoncés n'est jamais l'événement ou le sujet humain : les énoncés n'ont jamais comme objet que l'empreinte ou la trace de l'humain sous forme d'inscription.

Pour que l'inscription qui garde trace de l'événement puisse être objet de la recherche, c'est-à-dire qu'on puisse la décrire et l'interpréter, elle doit nécessairement posséder six caractéristiques de base : linéarisée et multilinéaire, analysable et isolable, identifiable et manipulable.

Toute inscription doit être linéarisée et multilinéaire, comme une partition musicale, en ce sens que l'espace et le temps doivent nécessairement y être inscrits. Dans sa fugacité, l'événement apparaît à la perception comme un ensemble global de divers éléments : plusieurs objets se meuvent en même temps dans le champ visuel, alors que plusieurs discours peuvent être tenus simultanément. Tout ce qui se passe en différents endroits et en même temps ne peut pas être retracé en un seul signe sur une seule ligne. L'enregistrement magnétoscopique peut garder une trace de tels événements, mais cette trace est globale, analogique, et, du point de vue de l'analyse, même si elle peut être reproduite, elle reste cacophonique. Pour que la trace globale soit analysable, l'enregistrement doit séparer les traces des éléments qui ont constitué l'événement. Chaque élément est représenté par une ligne, c'est-à-dire une suite de coordonnées, de symboles ou d'énoncés, et cette ligne est étirée en fonction d'un étalement du temps propice au repérage des changements les plus brefs. Une bonne analogie en est la partition musicale ou les enregistrements graphiques à plusieurs traceurs, comme ceux des électro-encéphalogrammes. L'inscription comporte donc une référence temporelle de base, c'est sa linéarité, qui permet de juxtaposer ou de superposer, afin de les comparer, les diverses lignes qui sont la trace particulière aux éléments qui, dans leur simultanéité, ont composé l'événement. L'inscription sera multilinéaire, car les traces de chacun des éléments qui se passent en même temps doivent pouvoir se lire avec leur simultanéité, comme lorsqu'on voit superposées plusieurs oscillations sur un écran. Cette linéarisation multilinéaire implique aussi une homogénéité du système d'expression des diverses inscriptions qui composent, analogiquement, les multiples lignes d'une partition d'orchestre.

Les inscriptions doivent être analysables. Le système d'inscription doit permettre le repérage des discontinuités, des alternances et des ruptures, sans quoi il est impossible de noter les régularités qui peuvent apparaître lorsqu'on suit les modulations dues au temps. Analogiquement, l'ampleur des variations et leur étalement doivent être suffisants pour qu'on puisse percevoir les variations sans devoir prendre une loupe. Mais de plus, les lignes qui composent ces inscriptions doivent être isolables : chacune doit pouvoir être lue indépendamment des autres, afin d'en repérer l'allure particulière, et chacune doit pouvoir être utilisée comme base pour comparer

les autres entre elles afin de percevoir les variations et les redondances spatiales. Autrement dit, il ne faudrait pas que le retrait d'une trace en arrière plan ou sa mise en avant plan empêche de voir et d'analyser les autres lignes.

Les inscriptions doivent encore être identifiables : le statut des éléments constitutifs de l'événement doit pouvoir être repéré. Entre autres choses, on doit pouvoir identifier qui ou quoi, dans l'échange, est sujet et objet, qui a l'initiative, qui clôture la séquence et quelle en est la référence, quelles sont les positions de chacun des partenaires et qui ou quel symbole joue le rôle de tiers témoin et régulateur, etc. En reprenant l'analogie de la partition musicale, le chef d'orchestre doit pouvoir identifier quelle ligne correspond à ce que jouera le premier violon par rapport à l'ensemble des cordes.

Enfin, toutes les inscriptions manifestant les caractéristiques précédentes ne seront porteuses de questions et de réponses que si elles sont manipulables, c'est-à-dire transformables. Une inscription n'est scientifiquement parlante que si elle permet des comparaisons, ce qui implique deux choses. D'abord, les éléments éloignées dans l'espace et dans le temps doivent être déplacés et rapprochés pour être comparés. Ensuite, les éléments exprimés dans des langages ou selon des systèmes différents doivent être traduits selon un code commun. Autrement dit, ce qui est intraduisible reste incomparable.

La trace de l'objet de recherche doit donc pouvoir être consignée dans un système qui permette de la désigner, de la dénommer, de la décrire : elle doit donc être écrite. Et comme la mise à l'épreuve des interprétations qui seront faites de la trace implique aussi la linéarité, l'analyse, l'identification et la manipulation, les énoncés produits à propos de la trace doivent être écrits eux aussi. Telles sont les exigences épistémologiques qui conduisent à écrire l'objet de la recherche.

3 Le caractère culturel et donc collectif de la recherche

À l'intersection des exigences épistémologiques et des nécessités stratégiques de la recherche, l'écriture joue aussi un rôle majeur, dans la mesure où l'activité de recherche n'a de sens que si elle s'adresse aux autres et leur parle.

Le chercheur, en tant que sujet énonciateur, parle d'un objet et non pas de lui-même. En recherche scientifique, l'énonciateur ne peut pas de lui-même parler sur lui-même. De manière autonome, c'est-à-dire en restant

sa propre référence, en étant lui-même le garant de sa propre loi, l'énonciateur ne peut produire ni inscription, ni question, ni discours de lui-même pour les autres, à moins d'être un dieu, de se prendre pour le nombril du monde et de s'imposer tyranniquement comme doctrine incontestable. L'auto-discours autonome n'a rien à voir avec la science. Pour que les énoncés puissent supporter des questions et des mises en questions, l'énonciateur ne peut les faire porter que sur d'autres. Lorsqu'un énonciateur produit des énoncés à propos de lui-même, ils ne peuvent être crédibles que s'ils constituent une réponse aux questions d'au moins un autre et se plient aux exigences du dialogue.

4 Les nécessités stratégiques de l'écriture

Il y a plus : tout énoncé de recherche qui n'est ni lu, ni commenté, ni contesté, c'est-à-dire qui n'est pas repris par d'autres et, en ce sens, qui ne participe pas au collectif de la recherche, n'existe pas en tant qu'énoncé scientifique : il n'est qu'un soliloque. D'où les nécessités stratégiques de l'écriture. L'énoncé scientifique n'existe que s'il intéresse d'autres scientifiques. Pour cela, le chercheur doit parler à d'autres, avec d'autres, à propos des autres. Et c'est la considération des autres qui est objet de stratégie.

4.1 *Écrire à d'autres avec leur style*

Il s'agit donc d'abord d'écrire à d'autres chercheurs. L'écrit a besoin de lecteurs et, pour les rejoindre, il lui faut un média. Une fois les lecteurs choisis, il faut repérer quelles revues ils lisent, quel est le style, le vocabulaire et les références de ces revues. Autrement dit, puisque les auteurs d'une revue ont aussi tendance à la lire, il s'agit de savoir comment les autres chercheurs, nos lecteurs anticipés, écrivent dans les revues qu'ils lisent. L'auteur court le risque d'être rejeté s'il veut imposer son style à ses lecteurs. C'est à lui d'adopter leur style afin d'être publié pour être lu dans les revues de ses lecteurs. Il s'agit donc de lire les autres, non seulement pour savoir ce qu'ils ont écrit, mais aussi pour savoir comment leur écrire.

4.2 *Écrire à propos des énoncés des autres*

Les lecteurs les plus assidus et les plus impatients de lire une critique sont les artistes à propos desquels la critique est écrite. La meilleure manière d'intéresser les autres à un écrit est d'écrire à propos de ce qu'ils ont écrit. La question est alors : qu'est-ce que les autres, ceux qui m'intéressent, ont

écrit qui puisse être contestable, ou du moins qui puisse être l'objet d'un commentaire intéressé ?

Or, sans nécessairement être voué à la controverse et à la polémique, l'énoncé dont on attend un commentaire est celui dont on n'est pas certain. Stratégiquement, il s'agit donc de repérer, dans les écrits des autres, quels sont leurs énoncés faibles, ceux qui ont été formulés avec audace, bluff ou imprudence, c'est-à-dire avant que les garanties et les preuves suffisantes ne soient accumulées. L'astuce, ou la force du nouvel énoncé, consistera à proposer une reformulation de leurs écrits qui comporte des précautions telles que cette reformulation ne pourra être perçue par les autres que comme une amélioration de l'énoncé initial. Dans bien des cas, on obtiendra cet effet en ajoutant des distinctions permettant de mieux identifier l'objet du discours et en précisant les conditions sous lesquelles l'énoncé contesté ou reformulé peut être maintenu.

4.3 Utiliser les écrits des autres

Une manière de se mettre à l'abri d'un effet boomerang et de donner du poids aux écrits à propos des autres (les contestés) consiste à trouver quelques écrits produits par d'autres (des alliés), sur lesquels on peut s'appuyer et derrière lesquels on peut se retrancher. Si ces alliés ont une bonne réputation, l'énoncé contestataire risque moins de subir l'accusation de n'être que l'écho d'un farfrelu ou la vengeance d'un marginal. En outre, plus les alliés d'un contestataire sont nombreux, moins celui qui est contesté pourra résister en isolant les écrits qui le critiquent. Latour² a fort bien montré l'importance des alliances dans la résistance d'un argument lors d'une controverse.

À cette phase de la stratégie, la question devient : quels sont, parmi ceux qui ont écrit, ceux sur lesquels on peut s'appuyer ? Autrement dit, quels pourraient être les alliés qui permettraient de convaincre les autres, les sceptiques et les contestés. Stratégiquement, l'auteur d'un énoncé est moins vulnérable s'il peut trouver chez d'autres, et chez de nombreux autres, des arguments qui fondent, qui supportent, qui illustrent ou qui rationalisent des éléments de son énoncé.

Lors de la recension des écrits, dont l'exhaustivité est impossible et dont la neutralité est illusoire, le chercheur se fixera donc trois objectifs :

- 1° S'imprégner du style d'écriture des revues que lisent ceux que l'on souhaite avoir comme lecteurs;

2 Latour B., *La science en action*. Paris : Éditions La Découverte, 1989.

- 2° Identifier l'un ou l'autre écrit contestable, ce qui permet d'engager une controverse et, par conséquent, d'avoir au moins un lecteur. Accessoirement, cette controverse contribuera à l'avancement de la discipline;
- 3° Asseoir l'argumentation de la controverse sur des alliances qui donnent du poids et qui empêchent les sceptiques de renvoyer la balle.

Académiquement ces objectifs sont généralement tus, sans doute parce que considérés comme trop cyniques et manquant de noblesse. Certains méthodologues préfèrent dire que la recension des écrits a pour but la construction du cadre conceptuel ou théorique qui permettra une formulation technique et opérationnelle du problème de recherche. L'attribution de cette seule finalité à la recension des écrits n'est acceptable que dans le cadre d'un travail académique comme peuvent l'être les mémoires et les thèses d'étudiants. Cependant, dans la mesure où la formation des jeunes chercheurs a des visées plus lointaines que l'obtention du diplôme, il est essentiel de faire comprendre aux étudiants que la formulation technique du problème ne sera efficace que si elle s'effectue en considérant les trois objectifs stratégiques qui viennent d'être énoncés. En effet, une problématique apparaît d'autant plus pertinente qu'elle participe à une controverse. Les choix méthodologiques semblent d'autant plus justifiés qu'ils s'appuient sur des alliances et que les protocoles d'instrumentation et d'analyse sont repris à d'autres chercheurs qui les ont raffinés et qui en ont établi, provisoirement, la validité.

4.4 Ne pas détourner les alliances ni se mettre hors de portée

Si l'on veut être certain de susciter un intérêt chez quelques lecteurs, les écrits porteront sur des énoncés où d'autres (les alliés potentiels) entretiennent déjà une controverse. L'intérêt augmentera si leurs alliances ne paraissent pas détournées au seul profit du chercheur, si ce dernier apporte des arguments supplémentaires qui renforcent les positions de ses alliés et s'il ne se place pas en dehors du débat, dans la position des membres d'un jury qui se contentent d'évaluer le travail des autres. Cette dernière condition est importante, car dans ce jeu d'alliances, la manière d'utiliser les arguments ne doit pas écarter l'auteur du collectif des chercheurs. En effet, se situer hors de la mêlée permet de se soustraire à la possibilité d'une controverse, mais amoindrit la crédibilité des énoncés qui ne sont plus considérés alors que comme des écrits exotiques.

4.5 *Distiller le doute*

La puissance de la controverse est plus forte si, plutôt que d'affirmer que l'erreur des énoncés contestés est évidente, les écrits se contentent de montrer que le contesté aurait dû introduire des nuances dans ses écrits, que ses affirmations doivent être modulées (par exemple, ses énoncés ne tiennent que dans certaines circonstances, et celles-ci sont rares...), qu'il conviendrait d'y introduire plus de relativité qu'il n'a bien voulu en mettre. Dans sa subtilité, la contestation sera plus aiguë encore si les modulations, que l'on regrette devoir imposer aux énoncés contestés, découlent de principes méthodologiques, de règles et de définitions que d'autres auteurs ont formulés et autour desquels un consensus peut être évoqué.

4.6 *S'inscrire dans une culture professionnelle*

Enfin, lorsque la recherche porte sur un des problèmes des sciences humaines et sociales (dans ce cas, l'éducation), il est de plus impérieux que les énoncés s'inscrivent dans une culture professionnelle. Sans cela, la dissémination des résultats de la recherche chez les professionnels ne se fera pas. Or, s'il peut être important, pour la promotion dans la carrière, d'être lu par les trois spécialistes au monde du problème qui passionne le chercheur, il est encore plus important d'être lu et utilisé par les professionnels dont les pratiques sont censées être fondées et instrumentées par la recherche. Dans un contexte où les agences gouvernementales réduisent les subventions à la recherche sociale, les meilleurs alliés des chercheurs seront les professionnels (et leurs clients), à condition qu'ils soient intéressés par la recherche, car ils constituent une masse politiquement bien plus importante que la caste des chercheurs.

S'inscrire dans une culture professionnelle signifie plusieurs choses. D'abord, il s'agit de respecter ses alliés potentiels en ne diminuant pas la pertinence de leurs énoncés sous le prétexte d'une faiblesse conceptuelle et en ne minimisant pas la portée de leurs questions parce qu'elles ne correspondent pas à la pureté théorique ou épistémologique souhaitée par les chercheurs. Les problèmes réels, ceux des professionnels du terrain, ne sont pas philosophiquement purs. En conséquence, l'incapacité des théories à aborder ces problèmes doit être acceptée comme une limite, comme une faiblesse inhérente à la théorisation face aux exigences et aux contraintes du monde réel. Les questions posées à partir d'une pratique professionnelle ne peuvent donc pas être déclarées comme mal formulées; ce sont les théories et les concepts qui sont trop éthérés.

4.7 *Proposer, non pas des solutions, mais des outils pour leurs solutions*

Enfin, que les lecteurs soient des alliés du chercheur ou des professionnels, utilisateurs potentiels des résultats de la recherche, il importe de leur montrer que ce n'est pas eux qui doivent faire un détour pour pouvoir utiliser les énoncés du chercheur, mais qu'au contraire, ces énoncés participeront à la solution de leurs problèmes. Par exemple, il faudrait présenter les énoncés produits par la recherche de telle sorte que les professionnels aient l'impression que ces énoncés peuvent même être utiles à la poursuite de leurs objectifs.

En somme, quelles que soient les alliances que les nécessités stratégiques imposent au chercheur, s'il veut que ses écrits soient lus, il faut qu'il montre à ses alliés qu'il connaît ce qu'ils savent, qu'il n'ignore pas les questions qu'ils se posent et que les énoncés produits de ses recherches sont utilisables pour leurs problèmes, non pas comme solution apportée de l'extérieur, mais comme des jalons, comme outils mis à leur disposition pour leurs propres solutions.

5 Les questions techniques

5.1 *La manière d'écrire : reformuler la pensée pour la communiquer*

Le premier problème d'écriture que rencontrent bien des chercheurs est la difficulté de passer de l'expression de leur pensée à la communication de leur pensée. Il est évident que les chercheurs vont utiliser l'écrit (manuscrit ou informatisé) pour consigner leurs notes de lectures, leurs analyses, leurs raisonnements. Ces notes leur servent à visualiser leurs impressions et leur pensée, elles leur permettent d'enchaîner les arguments d'un raisonnement et de le tester, elles leur servent aussi de mémoire temporaire : en tant que telles, ces notes sont adressées à leur propre conscience et elles sont impropres à être communiquées à des lecteurs. Trop souvent ces notes manquent de clarté, car elles sont le brouillon de la pensée et elles utilisent des termes avec des connotations personnelles, en ne leur accordant pas le sens admis par les lecteurs éventuels. Si ces notes sont indispensables, elle sont aussi source de malentendu quand elles ne sont pas complètement reformulées.

Lors de la reformulation, le chercheur sera attentif au mirage du jargon qui se manifeste autant chez le novice que chez l'expert. Le jeune chercheur

est souvent peu clair, parce qu'il ne maîtrise pas assez ce qu'il voudrait dire. Tant qu'il n'est pas capable d'expliquer, avec aise, ses arguments à un lecteur, il doit retravailler ses brouillons. Faire le savant en utilisant le jargon technique des spécialistes alors que cette langue n'est pas intégrée, conduit à un simulacre vite démasqué par les inconsistances qu'il produit. Le chercheur novice dans un domaine aura sans doute avantage à utiliser un vocabulaire non spécialisé, mais précis et concis, qui lui permet de communiquer sa pensée plutôt que de singer le jargon académique souvent ampoulé. Les vieux chercheurs sont parfois aussi peu clairs que les novices dans la mesure où, pressés de publier et de beaucoup publier, ils peuvent être tentés de cacher l'insuffisance de leur pensée par des jeux de langage, par la manipulation d'un vocabulaire ésotérique, par une syntaxe compliquée qui emboîte sans fin et sans nécessité les nuances et les circonlocutions qui perdent le lecteur. Ça fait savant, érudit, mais c'est tout. Même si cela peut séduire, au début, par un certain style, à la longue cela ne résiste pas à l'analyse.

Contrairement à ce que certains pensent, l'expression la plus directe et la plus précise est plus convaincante que le discours qui comporte trop de nuances, de mises en garde, de conditions, de modulations. Il convient aussi d'éviter les pseudo-néologismes qui ne font qu'exprimer la participation à une école de pensée au détriment de la clarté : la langue française attribue un sens plus clair aux mots élèves et étudiant qu'à «l'apprenant» ou au «s'éduquant». Une nouvelle terminologie n'est intéressante que dans la mesure où elle enlève des ambiguïtés; elle est à éviter quand elle introduit le flou et devient un fourre-tout. Un énoncé qui peut être interprété de dix manières différentes est un énoncé à proscrire; l'énoncé magique du psychanalyste parisien Lacan, «Ça parle», peut tout dire et, en conséquence, n'apporte rien de neuf.

5.2 *Savoir quel est le destinataire et s'adresser à lui*

Écrire une thèse, rédiger un article et préparer une demande de bourse ou de subvention sont des activités d'écriture très différentes. Avant de s'y attaquer, il faut savoir à qui s'adressent ces documents et ce que les différents types de lecteurs attendent de l'auteur.

Le lecteur d'une revue attend de l'auteur qu'il traite d'un sujet apparenté à l'objet général de la revue en utilisant le langage habituel des professionnels qui la lisent, y compris la structure de l'exposé, le genre d'illustration, le type d'argumentation et la mise en page.

Les administrateurs d'un organisme subventionneur, y compris les pairs qui siègent sur leurs comités, attendent que les auteurs respectent le formulaire, qu'ils répondent aux questions et utilisent les catégories qui y sont prévues, et cela avec clarté et concision, en se limitant au nombre de pages accordé et en fournissant toutes les informations demandées. Ils n'attendent pas que le demandeur étale tout son savoir et son expérience, ni qu'il présente une argumentation fouillée. Ils s'attendent à disposer de l'information leur permettant simplement d'évaluer si le projet est un projet pertinent et s'il a des chances d'aboutir.

Les membres du jury d'un mémoire ou d'une thèse s'attendent à ce que le candidat montre qu'il sait comment faire une recherche dans un domaine précis et qu'à l'occasion, il contribue, par un petit rien, à clarifier un des problèmes qui se posent dans ce domaine. Les membres du jury, ayant malheureusement souvent d'autres tâches plus urgentes à faire que d'évaluer les mémoires et les thèses, s'attendent à ce que cette démonstration de la maîtrise de la recherche se fasse avec clarté et concision.

5.3 Les phases de la recherche et le style d'écriture

Par rapport aux phases habituelles d'une recherche, l'exigence de clarté et de concision peut être précisée comme suit.

5.3.1 Le problème

Le problème, que les intellectuels ont pris l'habitude de compliquer pour en faire une problématique, sera décrit dans un langage et avec les perspectives contemporaines des professionnels du domaine. On part du problème tel qu'il se pose sur le terrain réel, et on laisse les discussions philosophiques sur son insertion dans la construction du monde aux autres. De ce point de vue, s'il est essentiel que le chercheur déclare sa position épistémologique, il est inutile qu'il alourdisse sa thèse d'un chapitre de discussion épistémologique et historico-critique, à moins que la thèse ne porte spécifiquement sur l'épistémologie.

5.3.2 Le cadre conceptuel ou théorique

La revue critique des écrits et sa synthèse, dénommée cadre théorique en recherche vérificative ou cadre conceptuel en recherche exploratoire, seront centrées sur les concepts et les notions nécessaires et suffisants à la mise en place des hypothèses et du dispositif de recherche.

Il est inutile de vouloir y écrire tout ce qu'on a lu : l'étalage d'érudition savante fournit surtout au jury l'occasion de voir les lacunes de l'étudiant. À trop écrire, l'étudiant multiplie les occasions de montrer ce qu'il n'a pas compris. De toute façon, à notre époque, il est impossible d'avoir lu tout ce qui traite d'un sujet et les membres d'un jury (bien constitué) ont probablement accumulé plus d'informations au cours de leur carrière que l'étudiant pendant ses études. Il vaut donc mieux ne pas se faire d'illusions et se limiter à la seule littérature nécessaire et suffisante au problème.

La recension critique des écrits (dite aussi revue de la littérature, ou revue de la documentation) se fait avantageusement, comme on l'a dit, dans une perspective stratégique. Il s'agit donc d'abord de trouver des énoncés contestables pour les critiquer et pour proposer des corrections. À cette fin, on cherchera particulièrement

- les auteurs qui ont tendance à produire des généralisations abusives;
- les prescriptions (méthodologiques et professionnelles) inopérantes, c.-à-d. ou bien inapplicables dans certains contextes ou bien qui ne fournissent pas ce qu'elles sont censées produire;
- les exceptions qui infirment les énoncés et qui limitent leur portée;
- les conditions d'application trop larges, faussement précises, qui permettent de faire bien d'autres choses;
- les critères flous, tels que les indices d'une catégorie d'événements qui recouvrent aussi d'autres catégories d'événements.

Par ailleurs la production d'énoncés alternatifs à ceux que l'on conteste se fait aisément

- en formulant des distinctions que les autres n'ont pas faites, ce qui permet de proposer de nouvelles relations entre événements ou entre concepts;
- en précisant les modalités d'application des concepts et des prescriptions, ce qui conduit à clarifier les critères qui permettront de reconnaître une trace comme l'indice incontestable d'un événement.

5.3.3 La chronique

La chronique de la recherche décrivant la méthodologie réellement suivie (ce qu'on a fait et non pas ce qu'on aurait dû ou voulu faire) devra, sans se perdre dans les anecdotes, être la plus explicite, la plus claire, la moins ambiguë possible par rapport

- aux présupposés et aux orientations théoriques privilégiées;
- aux contraintes qu'on a rencontrées, à celles qu'on a ignorées et à celles qu'on a respectées;
- aux décisions successives (changements d'outils, de stratégies) qui ont été prises et à leurs raisons;
- aux procédures concrètes qui ont été suivies tout au long de la constitution, de l'analyse, du traitement et de l'interprétation des données.

La clarté de la chronique devrait être telle que les lecteurs puissent admettre sans réserve les énoncés sur les traces produites par l'instrumentation, sans que des nuances puissent y être posées après coup. Pour cela, il faut que

- le statut des traces (les inscriptions ou le matériel) recueillies soit le plus clair, le plus transparent possible, c'est-à-dire tel qu'il ne faille pas, après coup, apporter des nuances et préciser des modalités pour limiter la signification et les interprétations;
- les analyses faites sur les traces soient compréhensibles sans qu'il faille y apporter des restrictions;
- les critères permettant les interprétations soient évidents, clairs, parlants, permettant de bien distinguer ce qui est considéré comme indice des événements cherchés de ce qui ne l'est pas.

6 Conclusion

Enfin, si le problème est clairement délimité, si le cadre théorique ou conceptuel est précis, si la chronique est transparente, si les interprétations sont explicites et consistantes, on pourra (idéalement) en arriver, dans les conclusions, à montrer que la question posée a été résolue de telle manière que la solution peut servir d'outil non contestable à d'autres pour poser et résoudre une autre question. Alors, notre recherche, sans faire de révolution, aura contribué à l'accroissement du capital culturel de notre société.