

Université de Montréal

Évaluation de la satisfaction des aides techniques en
réadaptation

par

Louise Demers

École de réadaptation

Faculté de médecine

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

Mai, 1999

© Louise Demers, 1999



Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

Évaluation de la satisfaction des aides techniques en
réadaptation

présentée par:

Louise Demers

a été évaluée par un jury composé des personnes suivantes:

Président du jury:	Raynald Pineault
Directrice de recherche:	Rhoda Weiss-Lambrou
Codirectrice de recherche:	Bernadette Ska
Membre du jury:	Elisabeth Dutil
Examineur externe:	Jeffrey Jutai
Représentant du doyen:	

Thèse acceptée le: _____

SOMMAIRE

Dans le domaine de la technologie en réadaptation, une aide technique désigne un objet spécifiquement conçu pour faciliter l'accomplissement des activités quotidiennes, promouvoir une plus grande autonomie et procurer une meilleure qualité de vie aux personnes présentant une déficience, une incapacité ou une situation de handicap. En dépit d'une forte croissance au sein de la population, encore peu de données permettent de déterminer dans quelle mesure les aides techniques répondent aux objectifs pour lesquels les usagers les utilisent.

Dans le cadre de l'évaluation des effets et conséquences d'aides techniques, les résultats de plusieurs recherches confirment l'importance de la satisfaction de l'utilisateur, parmi un ensemble d'autres indicateurs d'efficacité tels que l'indépendance fonctionnelle, la qualité de vie et les coûts. Cependant, la notion de satisfaction manque d'ancrage théorique et ses mécanismes d'opération sont encore peu connus. Dans le domaine de la réadaptation, cette lacune conceptuelle se répercute sur la rigueur des méthodologies employées pour mesurer le concept de satisfaction. À cet effet, il n'existe pas de moyen éprouvé pour recueillir la satisfaction de l'utilisateur d'aides techniques. Cette thèse a pour objet l'étude des valeurs métrologiques d'un instrument d'évaluation de la satisfaction envers les aides techniques utilisées en réadaptation. Ainsi, pour la première fois, la question de la fidélité et de la validité de l'*ÉSAT (Évaluation de la Satisfaction envers une Aide Technique)* a été étudiée de façon systématique.

Pour y arriver, trois étapes principales ont été réalisées. Lors d'une première étape, une étude d'envergure internationale sur la validité de l'*ÉSAT* a été entreprise, afin de vérifier si le contenu de l'instrument reflète bien le concept

mesuré. À cette fin, une douzaine d'experts provenant du Canada, des États-Unis et des Pays-Bas ont complété un questionnaire d'évaluation portant sur la pertinence des énoncés compris dans l'ÉSAT de même que sur ses procédures d'administration. Les résultats obtenus confirment que toutes les facettes importantes de la satisfaction envers les aides techniques sont correctement représentées. Cependant, des améliorations aux procédures d'administration et à la formulation des items pris individuellement sont souhaitées.

Deuxièmement, la fidélité de l'ÉSAT a été vérifiée, du point de vue de la stabilité test-retest et de la reproductibilité interjuges. Pour réaliser cette étude, 139 sujets ont été évalués à deux reprises à l'intérieur d'un intervalle de 7 à 11 jours. Pour un premier groupe de sujets ($n=85$), les deux évaluations ont été administrées par le même juge alors que, pour un second groupe ($n=54$), elles ont été conduites par des juges différents. Les coefficients de fidélité obtenus ont révélé que, pour la grande majorité des items, l'évaluation de la satisfaction génère des résultats stables (kappa pondéré variant de 0.42 à 0.74) et reproductibles (kappa pondéré variant de 0.30 à 0.66). Par contre, il en va autrement pour l'évaluation de l'importance attribuée aux énoncés, qui s'avère ni stable, ni reproductible (kappa moyen de 0.36). Ces derniers résultats s'expliquent par une très faible dispersion des données autour des moyennes, remettant en question la pertinence d'inclure cette partie d'évaluation au sein de l'instrument de mesure.

Pour la troisième étape de cette thèse, une analyse approfondie des énoncés compris dans l'ÉSAT a été réalisée, à partir de critères de rendement individuel (degré d'applicabilité, validité de contenu, validité de critère, contribution à l'homogénéité interne, stabilité test-retest et sensibilité) de même que collectif

(homogénéité interne et validité factorielle). Les résultats ont permis de réduire le nombre d'énoncés de moitié, tout en optimisant les propriétés de fidélité et de validité de l'ensemble. En outre, deux dimensions de satisfaction envers les aides techniques ont été mises en évidence. La première, la "technologie" comprend huit items alors que la deuxième, les "services", regroupe quatre items. Cette structure bifactorielle a été obtenue avec deux échantillons de sujets, de sorte que le modèle est considéré comme stable et valide.

En conclusion, cette thèse tient sa valeur du fait que l'ÉSAT permettra d'exercer une meilleure évaluation des effets et conséquences d'aides techniques attribuées aux personnes présentant une déficience, une incapacité ou une situation de handicap. En poursuivant la démarche entreprise avec l'élaboration d'une version révisée de cet instrument, des retombées positives sont anticipées relativement à l'attribution, au suivi, à la recherche et au développement d'aides techniques répondant mieux aux besoins des usagers.

TABLE DES MATIÈRES

Sommaire	iii
Liste des tableaux	viii
Liste des figures	ix
Liste des sigles et abréviations	x
Remerciements	xi
Dédicace.....	xiii
CHAPITRE 1: Introduction - - - - -	1
1. Évaluation des effets d'aides techniques.....	4
Définitions	4
Études théoriques	6
Instruments de mesure	10
2. Concept de satisfaction en réadaptation.....	13
Définitions	13
Études théoriques	14
Études empiriques	17
Conclusions	22
3. Propriétés métrologiques requises des instruments de mesure....	24
Fidélité	24
Validité	25
4. Travaux préliminaires.....	28
Développement de l'ÉSAT	28
Description de l'ÉSAT	29
5. Position du problème et objectifs généraux	32

CHAPITRE 2: An international content validation of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) - - - - -	35
Abstract.....	37
Introduction	38
Description of QUEST.....	39
QUEST as an assessment tool for occupational therapists	42
Methodology.....	44
Results	46
QUEST content coverage and item relevance.....	48
Administration procedures	52
Discussion	54
Conclusion.....	59
Acknowledgments.....	61
References.....	62
CHAPITRE 3: Stability and reproducibility of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) - - - - -	65
Abstract.....	67
Background	68
Methods.....	73
Subjects	73
Raters	76
Procedure	77
Statistical Analysis	77
Results	79
Stability and reproducibility of importance ratings.....	79
Stability and reproducibility of satisfaction ratings	83
Discussion	87

Acknowledgments	91
Références	92
CHAPITRE 4: Item analysis of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) - - - - -	95
Abstract	97
Introduction	98
Methods	102
Measurement properties and related parameters	102
Data sets	103
Item selection plan	106
Part one: Decision criteria	106
Part two: Decision making.....	109
Cross-validation study.....	110
Results	111
Items performance	111
Selection of the best performing items.....	114
Cross-validation	117
Discussion	119
Acknowledgments	124
References	125
CHAPITRE 5: Discussion générale et conclusions - - - - -	129
1. L'évaluation des effets d'aides techniques	131
Études théoriques	131
Instruments de mesure	134
2. Le concept de satisfaction en réadaptation	136
Définitions	136
Études théoriques	138

Études empiriques	140
3. Éléments de réponse et nouvelles questions	142
Références	145
Appendice A: Formulaire d'évaluation de l'ÉSAT	xvi
Appendice B: Certificat d'éthique.....	xvii

LISTE DES TABLEAUX

Chapitre 1

Tableau I: Dimensions de l'évaluation des effets et conséquences d'aides techniques en fonction des types d'utilisateurs de données	9
---	---

Chapitre 2

Table I: QUEST items and definitions	40
Table II: International content experts of QUEST	47
Table III: Percentage of agreement on the relative importance of items	49

Chapitre 3

Table I: Subjects, devices and environment characteristics of the two samples	75
Table II: Descriptive statistics of QUEST importance and satisfaction ratings	80
Table III: Coefficients of stability and reproducibility of 24 QUEST items: Importance ratings	82
Table IV: Coefficients of stability and reproducibility of 20 QUEST Items: Satisfaction ratings	84

Chapitre 4

Table I: Decision criteria for item selection.....	107
Table II: Measurement properties of the 24 QUEST items.....	112
Table III: Results of factor analysis (after orthogonal rotation) of 12 satisfaction items, using the Montreal sample (n=150)	116
Table IV Results of factor analysis (after orthogonal rotation) of 12 satisfaction items, using the Dutch sample (n=253)	118

LISTE DES FIGURES

Chapitre 1

Figure 1: Modélisation de la satisfaction en réadaptation 23

Figure 2: Étapes de développement de l'ÉSAT 30

Chapitre 3

Figure 1: Frequency distributions of items per category of agreement (n=6) 86

LISTE DES SIGLES ET ABBRÉVIATIONS

CERTAIN	<i>Cost-Effectiveness of Rehabilitation Technology through Appropriate Indicators</i>
D-QUEST	<i>Dutch version of the Quebec User Evaluation of Satisfaction with assistive Technology</i>
ÉSAT	Évaluation de la satisfaction envers une aide technique
IRV	<i>Institute for Rehabilitation Research</i>
KMO	<i>Kaiser-Meyer-Olkin measure of sampling adequacy</i>
MIF	Mesure de l'indépendance fonctionnelle
MMSE	<i>Mini-Mental State Examination</i>
OT	<i>Occupational Therapy</i>
PAF	<i>Principal Axis Factoring</i>
PCA	<i>Principal Components Analysis</i>
PIADS	<i>Psychosocial Impact of Assistive Devices</i>
QUEST	<i>Quebec User Evaluation of Satisfaction with assistive Technology</i>
RESNA	<i>Rehabilitation Engineering and Assistive Technology Society of North America</i>
SAT PRO	Satisfaction face à la prothèse
SPSS	<i>Statistical Package for Social Science</i>
TIDE	<i>Telematics applications for the Integration of the Disabled and the Elderly</i>
TNO-PG	<i>Dutch organization for applied scientific research - Prevention and health</i>
WFOT	<i>World Federation of Occupational Therapists</i>

REMERCIEMENTS

Lorsque j'ai entrepris, après plusieurs années de travail clinique, cette aventure extraordinaire de formation en recherche, je ne m'attendais pas à trouver autour de moi tant de support et de confiance. Cet appui m'a été offert sous diverses formes et à divers moments. C'est avec beaucoup de sincérité que je souhaite remercier ceux et celles sans qui je n'aurais pu mener ce projet à terme.

Tout d'abord, je tiens à exprimer ma vive reconnaissance envers mes directrices, Rhoda Weiss-Lambrou et Bernadette Ska, qui m'accompagnent depuis le début de ma formation graduée. Merci à Rhoda, qui m'a fait faire le pont entre la clinique et la recherche, pour sa confiance, sa vision grand angulaire de nos projets, sa grande disponibilité, sa motivation toujours renouvelée et son amitié. Sans elle, les travaux compris dans cette thèse n'auraient pas eu le même rayonnement international. Merci à Bernadette, mon ange gardien, dont la rigueur scientifique n'a d'égal que ses qualités humaines extraordinaires. Ses conseils et encouragements sont toujours venus au bon moment, me permettant de traverser avec plus de sérénité les périodes d'incertitude.

Merci aux membres de mon comité aviseur, Elisabeth Dutil et Raynald Pineault. Ils sont des sources inépuisables d'informations et de support. Merci à Yves Joannette, directeur du Centre de recherche de l'Institut universitaire de gériatrie de Montréal, qui a su allumer en moi l'étincelle de la recherche et de temps en temps en raviver la flamme. Merci à Daniel Bourbonnais, responsable des études supérieures à l'École de réadaptation, qui m'a soutenu efficacement tout au long de ma démarche. Merci à l'ensemble des

sujets qui nous ont si gentiment accueilli chez eux de même qu’au personnel du Centre de réadaptation Lucie Bruneau, de l’Institut de réadaptation de Montréal et du Centre de réadaptation Constance Lethbridge. Je remercie tout particulièrement Jacques Corbeil, pour son accueil chaleureux au sein du programme d’aides techniques qu’il dirige.

Merci à Luc DeWitte et à Roelof Wessels, nos collaborateurs des Pays-Bas de même qu’aux experts québécois, canadiens et américains qui ont testé l’ÉSAT dans leur milieu. L’intérêt qu’ils ont manifesté pour nos travaux m’ont donné des ailes. Je tiens aussi à remercier celles qui m’ont assistée pour la collecte des données, Camille, Michèle, Brigitte et Annick. J’ai apprécié leur disponibilité et la grande qualité de leur travail.

Merci aux organismes qui ont suffisamment cru en moi pour m’accorder des bourses et des subventions. Dans cette optique, je tiens à mentionner le Fonds de la recherche en santé du Québec, la Fondation canadienne d’ergothérapie et l’Université de Montréal. Merci aussi à Romedi Passini et au Fonds pour la formation des chercheurs et l’aide à la recherche de m’avoir permis de terminer la thèse sans problème.

Merci à Bertrand, mon mari, à qui je dois plus que les mots peuvent exprimer. En dépit de tous les risques et de toutes les contraintes, il a su m’encourager, me soutenir et faire en sorte que je maintienne le cap sur les objectifs qui importent dans ma vie. Merci à ma mère, à Laurent et à mes enfants, Thierry, Geoffroy et Micaëla, pour les sourires, les petites et les grandes attentions. Merci à mes amis, Francine, Brigitte, Caroline, Jo, Guy, Richard et Cathy, pour les encouragements et les moments de détente. Grâce à eux tous, je sais que je peux constamment aller plus loin.

À Bertrand, Thierry, Geoffroy et Micaëla,

Chapitre 1
Introduction

L'évaluation des effets et conséquences des interventions rendues dans le contexte des services de santé constitue une démarche en forte croissance. En effet, les administrateurs, bailleurs de fonds et la population exigent de plus en plus ouvertement des preuves d'efficacité et de rentabilité. Alors que les ressources financières restent insuffisantes pour combler la demande actuelle, le système de santé québécois doit composer avec une augmentation des besoins de sa clientèle. Ailleurs en occident, des problématiques similaires sont rapportées. Il devient par conséquent impérieux d'exercer un suivi relativement à la qualité, l'efficacité, l'utilisation et la satisfaction envers les interventions rendues.

Parallèlement, on assiste à la création et à l'expansion de ressources de réadaptation visant à limiter les situations de handicap et à favoriser le maintien à domicile des personnes présentant des incapacités. Parmi celles-ci, les aides techniques se diversifient et deviennent de plus en plus sophistiquées, produisant une augmentation de leur présence au sein de la population. En dépit de cette croissance, encore peu de données permettent de déterminer dans quelle mesure les aides techniques répondent aux objectifs pour lesquels les usagers les utilisent. À cet effet, il n'existe pas à l'heure actuelle de moyen systématique de recueillir la satisfaction des usagers que l'on souhaite associer à la démarche évaluative.

La présente thèse a pour objet l'étude des valeurs métrologiques d'un instrument d'évaluation de la satisfaction envers les aides techniques utilisées en réadaptation. L'instrument étudié est désigné par l'acronyme ÉSAT, pour Évaluation de la Satisfaction envers une Aide Technique. Dans les chapitres qui suivent, trois principaux objectifs s'inscrivent dans cette démarche

méthodologique sont abordés. Le premier objectif consiste à établir la validité de contenu de l'instrument, condition considérée comme préalable à l'étude des autres propriétés de mesure. Le deuxième objectif vise à cerner la stabilité test-retest et la reproductibilité interjuges des items qui composent l'ÉSAT. Ces deux caractéristiques des items sont considérées comme les indices de la fidélité de l'ÉSAT. Une fois les deux premiers objectifs rencontrés, il devient possible d'aborder le troisième objectif. Se basant sur les résultats obtenus, celui-ci vise (1) à analyser en profondeur chacun des items constituant l'instrument et (2) à vérifier s'il existe un sous-ensemble d'items présentant des propriétés de mesure optimales. À travers ces étapes successives de développement de l'ÉSAT, cette thèse contribuera à la mise au point d'une échelle de satisfaction permettant de documenter les impacts des aides techniques utilisées en réadaptation.

Avant d'aborder de façon directe les trois objectifs exposés brièvement plus haut, un cadre de référence plus général sera offert, afin de permettre au lecteur de mieux situer cette thèse au sein des connaissances déjà disponibles. Le développement d'un instrument d'évaluation de la satisfaction envers les aides techniques est une démarche qui est fortement encouragée dans le domaine de la réadaptation. Cependant, au moment d'entreprendre cette thèse, il n'existait encore à notre connaissance aucune étude réalisée à cet effet. Par conséquent, des notions empruntées à différents champs d'études ont dû être agencées pour une première fois ensemble en vue de soutenir les travaux de recherche prévus.

Dans un premier temps, les études portant sur l'évaluation des effets et conséquences de l'utilisation d'aides techniques sur la personne et sa qualité de vie seront revues. L'analyse de ces écrits permettra de cerner les besoins pour

des instruments tel que l'ÉSAT et de situer celui-ci au sein d'autres indicateurs d'efficacité susceptibles d'être utilisés concurremment. Dans un deuxième temps, une analyse des études théoriques et empiriques portant sur le concept de satisfaction sera présentée. Ces notions seront analysées et interprétées de façon à pouvoir s'appliquer aux aides techniques, ce qui n'avait pas été fait jusqu'à présent. Si les écrits portant sur la satisfaction comportent encore bien des zones grises, il en va autrement du domaine de la métrologie. Les propriétés de mesure requises pour utiliser avec discernement un instrument de mesure dans le domaine de la santé et en psychologie ont été amplement décrites. Les notions pertinentes à la réalisation de cette thèse seront résumées dans la troisième partie de la revue de la littérature.

Avant de passer à la présentation des articles, les travaux sur l'ÉSAT qui ont été réalisés avant d'entreprendre cette thèse seront succinctement présentés. Cette mise en contexte vise à bien situer le lecteur par rapport à l'historique et à l'état d'avancement de l'instrument, au moment de la mise en route de la thèse. L'ensemble des notions extraites de ce chapitre d'introduction imprégneront les devis méthodologiques et les interprétations des résultats présentés par la suite.

1. Évaluation des effets d'aides techniques

Définitions

Dans le domaine de la réadaptation, une aide technique désigne un objet spécialement conçu et prescrit pour compenser les déficiences et les incapacités des personnes handicapées de même que pour favoriser leur intégration sociale, scolaire et professionnelle (Conseil consultatif sur les aides technologiques, 1994). Depuis quelques années, les aides techniques

connaissent un essor important, en raison des changements qui s'opèrent dans le système de santé, tels que le virage ambulatoire et l'augmentation des besoins liés au vieillissement de la population. Cette croissance s'exprime par rapport aux types d'aides techniques disponibles alors que les champs d'application se multiplient et incluent des activités aussi variées que le positionnement, les soins personnels, les déplacements, la communication, les activités domestiques, le travail et l'aménagement domiciliaire. Elle se conçoit aussi en termes de coûts, lesquels sont estimés à environ 300 millions de dollars annuellement, uniquement pour la province de Québec (Office des personnes handicapées du Québec, 1993).

Dans ce contexte, il est devenu nécessaire pour les acteurs du domaine des aides techniques de rendre compte de leurs actions à la population et à ses représentants (Oortwijn, Boxsel & Beekum, 1997; Scherer, 1996a; Scherer & Vitaliti, 1997; Smith, 1996a; Trachtman, 1994; Warren, 1993). Selon plusieurs auteurs, il s'agit même d'une priorité dont l'importance va encore s'accroître dans le futur (Cook & Hussey, 1995; DeRuyter, 1995; Lawlor, 1991). En langue anglaise, les termes consacrés de "outcomes measures" sont utilisés pour désigner les évaluations des effets et conséquences découlant d'interventions de diverses natures. Les "outcomes measures" sont généralement définies de façon globale comme les effets, résultats et conséquences mesurables de l'intervention (Hammel, 1996; Keith, 1995; Lawlor, 1991; Scherer, 1996a, Trachtman, 1994), ce qui laisse place à une grande variété d'indicateurs. D'après Minkel (1996), le mot "outcome" réfère plus spécifiquement aux résultats de l'application d'un ensemble de ressources, incluant les aides techniques et les services associés, par rapport aux objectifs d'indépendance

fonctionnelle fixés conjointement par la personne, son entourage de même que les intervenants impliqués. Dans cette définition, les impacts des aides techniques sur la réduction des incapacités est explicite.

Plusieurs éléments contribuent à la complexité de l'évaluation des effets et conséquences d'aides techniques. D'abord, il est difficile d'isoler leur impact par rapport aux autres interventions inhérentes à la réadaptation d'une personne (Smith, 1996b). Ensuite, les résultats ne sont pas statiques mais en mouvance du fait, par exemple, de la fluctuation des incapacités, de l'apparition de technologies susceptibles de rendre obsolètes les précédentes, ou encore des changements dans les besoins de la personne (Hammel, 1996; Smith, 1996b). Minkel (1996) et Lawlor (1991) ajoutent que les résultats doivent être appréhendés à l'intérieur des limites des ressources disponibles, ce qui contribue aussi à la complexité du phénomène.

Pour recueillir des données sur les impacts d'aides techniques, deux éléments doivent être pris en considération. D'abord, il est primordial de préciser la nature des effets recherchés. Ensuite, il faut disposer d'outils de mesure appropriés. Ces deux aspects de l'évaluation des aides techniques sont abordés dans les sections qui suivent.

Études théoriques

Quelques-uns des travaux recensés sur l'évaluation des aides techniques se situent dans une approche de recherche et de développement de produits (Feeney, 1987; Jacques, Ryan, Naumann, Milner & Cleghorn, 1994; Ryan, Rigby & From, 1996). D'autres, la majorité, se situent dans une approche de pratique clinique et visent à aider les cliniciens à attribuer l'aide technique la plus

appropriée aux besoins de son client (Bain, 1997; Cook & Hussey, 1995; Scherer, 1996b; Schoech, Cavalier & Hoover, 1993;). Ces auteurs proposent des modèles conceptuels pour encadrer la démarche d'attribution tels que les modèles "*Bain Assistive Technology System (BATS)*" (Bain, 1997), "*Matching Person and Technology (MPT)*" (Scherer, 1996b) de même que "*Human Activity Assistive Technology (HAAT)*" (Cook & Hussey, 1995). Tous ces modèles ont en commun de souligner l'importance de la participation de l'utilisateur et de son entourage, que ce soit pour définir les besoins à combler, préciser les objectifs à poursuivre ou développer de nouveaux produits. Cependant, ils ne permettent pas de cerner la nature des effets recherchés dans une perspective d'évaluation des résultats.

Quelques ébauches de cadres conceptuels pour des évaluations de suivi sont proposées par Eblen (1992), Lawlor (1991) et Kohn, Mortola, et LeBlanc (1991). Ces auteurs établissent clairement une distinction entre l'évaluation des effets et l'évaluation du processus. Dans le premier cas, ils réfèrent au degré d'atteinte des buts visés, alors que, dans le deuxième cas, ils analysent comment l'aide technique influence la réalisation des activités de l'utilisateur. Pour rendre compte de l'efficacité d'aides techniques, les modèles proposés soulignent l'importance d'une part de la variable temporelle et, d'autre part, des priorités de l'utilisateur, puisque c'est l'interaction de ce dernier avec l'aide technique qui détermine le succès de l'intervention. D'après Eblen (1992), l'aide technique devrait ultimement être perçue par l'utilisateur comme une extension de sa personne. Comme indicateurs de résultats, Lawlor (1991) propose de considérer la sécurité de l'aide technique, la satisfaction des usagers, la satisfaction des intervenants de même que les coûts. Quant à Kohn et ses collaborateurs (1991),

ils s'intéressent au rendement du produit, à la performance de l'utilisateur de même qu'à son degré de satisfaction.

Un pas important dans la conceptualisation de l'évaluation des effets et conséquences d'aides techniques a été fait par DeRuyter (1995) qui a documenté leur caractère multidimensionnel. Considérant les paramètres les plus fréquemment cités dans la littérature à leur sujet, il délimite cinq dimensions à l'évaluation, soit (1) les résultats cliniques, (2) le rendement fonctionnel, (3) la qualité de vie, (4) la satisfaction de l'utilisateur et (5) les coûts. Les liens entre ces dimensions ne sont toutefois pas explicités. Une autre contribution importante est apportée par Smith (1996b), qui s'intéresse pour sa part aux fins d'utilisation des données évaluatives. Ses travaux l'amènent à proposer quatre types d'utilisateurs, soit (a) la société, dont les élus déterminent le soutien des programmes, (b) le programme de services, du point de vue de son contenu et de son application, (c) la personne, par rapport à l'impact sur ses capacités et ses habitudes de vie, et (d) les chercheurs. Compte tenu de cette diversité d'utilisation, les évaluations les plus pertinentes sont susceptibles de varier et peuvent inclure, toujours d'après Smith (1996b), le rendement de l'intervention, le rendement de l'utilisateur, les ressources et le support de l'environnement, les coûts ou encore une combinaison de ces aspects.

Dans une publication récente, Scherer et Vitaliti (1997) reprennent les dimensions de l'évaluation définies par DeRuyter (1995) et analysent leur pertinence pour diverses catégories d'utilisateurs. Cette grille d'analyse est présentée au tableau I. On peut remarquer que, mis à part "les chercheurs", les catégories proposées par Scherer et Vitaliti (1997) corroborent les types d'utilisateurs proposés par Smith (1996b). De plus, certaines dimensions

Table I: Dimensions de l'évaluation des effets et conséquences d'aides techniques en fonction des types d'utilisateurs de données *

Utilisateurs	Résultats cliniques	Statut fonctionnel	Qualité de vie	Satisfaction de l'utilisateur	Coûts
Administrateurs	non	oui	non	oui	oui
Usagers	non	oui	oui	oui	non
Cliniciens	oui	oui	oui	oui	non
Payeurs	oui	oui	non	oui	oui

* Traduit de Scherer & Vitaliti, 1997 par l'auteur (LD).

apparaissent être pertinentes pour l'ensemble des utilisateurs de données évaluatives. C'est notamment le cas pour la satisfaction de l'utilisateur.

Dans une perspective essentiellement socio-économique, une structure matricielle à trois dimensions visant à guider le développement d'instruments de mesure et de stratégies de recherche a été élaborée dans le cadre du projet européen "CERTAIN" ("*Cost-Effectiveness of Rehabilitation Technology through Appropriate Indicators*") (Oortwijn et al., 1997). Dans ce projet, les technologies visées englobent les aides techniques sans toutefois s'y limiter. Les axes considérés se rapportent (1) aux catégories d'utilisateurs de données: recherche, industrie, hôpital/intervenants et gouvernement/assureurs, (2) aux étapes du cycle technologique: développement/expérimentation, introduction sur le marché/acceptation, diffusion et pratique généralisée, ainsi que (3) aux types d'applications technologiques: aides techniques, interventions en milieu de travail, etc. La perspective de l'utilisateur n'est pas explicitement comprise dans la structure évaluative. Par conséquent, elle ne semble pas être pas considérée

comme indispensable à l'évaluation, ce qui distingue nettement ce cadre de tous les précédents.

Instruments de mesure

À partir du tableau I, on peut concevoir une variété d'instruments de mesures susceptibles de documenter les effets et conséquences d'aides techniques. Premièrement, les résultats cliniques s'avèrent relativement faciles à produire, à partir des indices clinimétriques normalement utilisés (amplitude articulaire, cicatrisation, etc.). Cependant, leur validité peut être mise en doute du fait que plusieurs autres variables non reliées aux aides techniques influencent les résultats obtenus (Minkel, 1996). Pour mesurer la deuxième dimension qui est le statut fonctionnel, divers indicateurs basés sur l'observation des tâches réalisées par la personne peuvent être utilisés. Le statut fonctionnel réfère à la performance dans les activités de vie humaines courantes, incluant les soins personnels, la mobilité, le travail, etc. (Liang et Jette, 1981). À cet effet, l'un des instruments les plus connus est la "Mesure de l'indépendance fonctionnelle" (MIF) (Uniform Data System for Medical Rehabilitation, 1995). Cette évaluation est critiquée à cause de son échelle de mesure, conçue de telle sorte qu'une personne utilisant des aides techniques n'est jamais considérée comme entièrement indépendante. Les tenants du "*Independence Living Model*" (Batavia, DeJong & McKnew, 1991) contestent cette vision inspirée du modèle médical et centrée sur les déficiences des individus. À leur avis, les situations de handicap sont attribuables aux obstacles créés par des environnements mal adaptés aux besoins des populations. Une ou des aides techniques peuvent donc supprimer la situation de handicap et rendre la personne indépendante. Par ailleurs, la fidélité et la validité de la MIF ont été étudiées auprès de

plusieurs clientèles types de la réadaptation (Brosseau, Philippe, Dutil & Boulanger, 1996; Demers & Giroux, 1997). Malgré un certain questionnement sur ses bases théoriques, l'ensemble des résultats s'avèrent positifs et tendent à soutenir l'utilisation de l'instrument (Brosseau et al., 1996).

L'instrument "OT Fact" (Smith, 1995) tente de corriger cette lacune de la MIF en permettant de comparer le rendement de la personne avec et sans aide(s) technique(s). Entièrement informatisé, cet instrument permet de détailler les aspects de la performance pour lesquels la personne requiert de l'aide, à travers des niveaux successifs d'embranchements de questions. Les quatre dimensions représentées sont l'intégration des rôles, le statut fonctionnel, l'intégration des habiletés de même que les composantes de la performance (Smith, 1995). Les qualités métrologiques du "OT Fact" n'ont pas été publiées.

La troisième catégorie d'évaluations concerne les impacts des aides techniques sur la qualité de vie. Plusieurs instruments de mesure peuvent être utilisés pour cerner ce concept en réadaptation (Day & Jankey, 1996). Cependant, Day et Jutai (1996) considèrent qu'un instrument dédié spécifiquement aux aides techniques est requis pour cerner plus précisément leurs impacts sur la qualité de vie. Dans ce but, ils ont développé un instrument intitulé "*Psychosocial Impact of Assistive Devices*" (PIADS) (Day & Jutai, 1996). Des analyses par composantes principales ont permis de construire trois sous-échelles reliées respectivement au sentiment de compétence, à l'adaptabilité et à l'estime de soi. Les propriétés de mesures de l'instrument sont à l'étude.

En ce qui a trait à la troisième catégorie de mesures, la satisfaction des usagers, la pénurie d'instruments est quasi complète. Cette conclusion a motivé la

création de “l'Évaluation de la Satisfaction envers une Aide Technique” (ÉSAT) (Demers, Weiss-Lambrou & Ska, 1996), dont l'étude des caractéristiques métrologiques fait l'objet de cette thèse. Récemment, un autre instrument de mesure de la satisfaction: le “SAT PRO”, a été publié (Bilodeau, Hébert & Desrosiers, 1998). Une seule catégorie d'aide technique est visée, soit les prothèses des membres inférieurs. En termes de propriétés de mesure, la consistance interne, la stabilité test-retest et la validité de construit du “SAT PRO” ont été vérifiées et se sont avérées acceptables.

Les coûts associés aux aides techniques constituent la dernière des grandes catégories d'évaluations de leurs effets et conséquences. Cependant, un seul instrument a été recensé, développé dans le cadre du projet européen *CERTAIN*. Intitulé le “CERTAIN tool”, cet instrument de mesure coûts-résultats serait applicable à l'évaluation de cas individuels d'utilisation d'aides techniques (Andrich, Ferrario & Moi, 1997, 1998). Des études sont en cours pour raffiner cet instrument et évaluer ses propriétés de mesure.

Cette section de la revue de la littérature sur l'évaluation des effets et conséquences d'aides techniques permet de constater un essor récent de travaux dans le domaine. Au plan théorique, des balises se dessinent et, bien qu'il ne s'agisse pas à proprement parler de cadres théoriques, les travaux des chercheurs dans le domaine tendent à s'articuler autour de thèmes convergents. Au plan des instruments de mesure, le besoin pour des instruments standardisés, fidèles et valides ressort de façon claire, particulièrement en regard des dimensions qualité de vie, satisfaction des usagers et coûts. Le développement de l'ÉSAT apparaît entièrement justifié dans le contexte de l'état d'avancement des mesures des effets et conséquences d'aides techniques.

2. Concept de satisfaction en réadaptation

Définitions

Il est difficile d'analyser la satisfaction sans aborder, ne serait-ce que brièvement, la notion de qualité, puisque ces concepts sont étroitement reliés dans la littérature (Finlayson & Havixbeck, 1992; Heinemann, Bode, Cichowski & Kan, 1997; Kohn et al., 1991; Locker & Dunt, 1978; McMillan, 1987; Simon & Patrick, 1997; Ware, Davies-Avery & Stewart, 1978). Figure dominante dans le domaine de la recherche sur la qualité des soins, Donabedian (1980) définit la qualité comme l'équilibre entre les bénéfices et les préjudices découlant des soins de santé, tel que perçu par un client bien informé ou par son répondant. Autour de ce concept gravitent des notions telles que l'accessibilité, la continuité et la coordination des soins. La satisfaction du client, pour sa part, permet d'apprécier et de mesurer la qualité des soins et donc de porter un jugement sur les résultats obtenus. Cette caractéristique de la satisfaction la distingue des autres approches d'évaluation de la qualité qui ciblent, par exemple, la structure et les processus de soins en cours d'intervention. En somme, bien qu'ils soient complémentaires et fréquemment associés dans une démarche évaluative, les concepts qualité et satisfaction ne s'équivalent pas et ne doivent pas être confondus.

Selon Keith (1998), la notion de satisfaction en réadaptation réfère à une attitude vis-à-vis un service, des intervenants ou un état de santé. Elle englobe des composantes d'ordre affectif, générant des émotions positives ou négatives, et d'ordre cognitif, s'exprimant par rapport aux dimensions évaluées et aux méthodes d'évaluation. Une définition quelque peu différente est donnée par Simon et Patrick (1997) pour qui la composante affective de la satisfaction est

primordiale. Pour ces chercheurs, la satisfaction reflète un degré de contentement, de bien-être et de gratification que le client ressent suite à une intervention de réadaptation, perçue globalement ou par rapport à ses composantes les plus importantes.

Caractériser, spécifier et définir la satisfaction seraient plus simples si les mécanismes qui engendrent cette attitude et/ou émotion étaient mieux connus. Or jusqu'à présent, les déterminants de la satisfaction restent imprécis, voire même indéfinis, ce qui nuit considérablement à l'opérationnalisation du concept et la rend très arbitraire (Aharony & Strasser, 1993; Bond & Thomas, 1992; Carr-Hill, 1992; Fitzpatrick, 1991; Hall & Dornan, 1988). C'est quasi à l'unanimité que les chercheurs soulèvent l'épineux problème créé par l'absence d'un cadre théorique de la satisfaction (Aharony & Strasser, 1993; Carr-Hill, 1992; Hall & Dornan, 1988; Keith, 1998; Williams, 1994). En conséquence, il est juste de se demander s'il est possible de comprendre concrètement ce que l'on mesure. En effet, les quelques modèles de la satisfaction existants ne sont que faiblement supportés par les recherches empiriques. De plus, ils ne visent pas spécifiquement le domaine de la réadaptation, bien qu'ils apparaissent suffisamment génériques pour l'embrasser.

Études théoriques

À la fin des années 1970, Ware et ses collaborateurs (1978) de même que Locker and Dunt (1978) effectuent des travaux consistant à réviser les instruments de mesure de la satisfaction existants envers les services médicaux. À partir de ceux-ci, Ware et ses collaborateurs (1978) tentent de définir le concept de satisfaction en identifiant ses principales dimensions de

mesure. Leur principale contribution à l'avancement des connaissances consiste à proposer une taxinomie de la satisfaction fondée sur huit éléments, soit: l'art des soins, la qualité technique des soins, l'accessibilité/la commodité, les coûts, l'environnement physique, la disponibilité, la continuité et l'efficacité. Les auteurs assument la présence d'interactions entre les facettes de la satisfaction mais ne considèrent pas qu'elles sont redondantes entre elles. Cependant, d'autres chercheurs marquent leur désaccord par rapport à cet avis et estiment que, quelque soit la classification employée, les dimensions de la satisfaction restent généralement fortement corrélées entre elles (Haddad, 1991; Pascoe, 1983). Dans les années 1980 et 1990, les chercheurs réfèrent fréquemment à la taxinomie en huit classes de Ware et ses collaborateurs (1978). À travers celle-ci, tous préconisent une approche multidimensionnelle de la satisfaction qui se perpétue jusqu'à aujourd'hui. Au plan théorique toutefois, la signification profonde du concept de la satisfaction reste inexplorée et la nature des liens entre ses dimensions n'est pas abordée.

D'autres chercheurs se sont penchés sur les déterminants de la satisfaction. Ainsi, les travaux de Linder-Pelz (1982a, 1982b) portent sur l'élaboration d'une théorie psychologique de la satisfaction envers les services médicaux. Cette chercheuse propose un modèle qui intègre des éléments d'ordre affectif et cognitif, respectivement reliées aux valeurs et aux attentes du client. Dans ce modèle intitulé "*Expectancy-Value Model*", la satisfaction découle d'une combinaison entre la force des valeurs du client et son attitude envers chacune des dimensions du phénomène étudié. Pour vérifier la validité de son modèle, Linder-Pelz (1982b) définit et teste cinq hypothèses liées aux valeurs, aux

attentes et à la perception du client. Les résultats s'avèrent toutefois peu concluants et tendent à discréditer le modèle.

Deux autres types d'approches théoriques de la satisfaction sont toujours d'actualité. En langue anglaise, les approches du premier type sont désignées comme les "*Discrepancy Theories*" et les approches du deuxième type sont connues comme les "*Fulfillment Theories*" (Bramadat & Driedger, 1993; Carr-Hill, 1992; Pascoe, 1983; Ruggeri, 1994). Dans le contexte des "*Discrepancy Theories*", la satisfaction correspond à l'écart existant entre les effets d'un épisode de soins et le résultat tel qu'anticipé par le client. En accord avec cette définition, l'évaluation de la satisfaction est rendue opérationnelle en cernant les perceptions du client en regard de ses valeurs initiales, de ses attentes et de son expérience des soins. Ces données permettent d'extrapoler ensuite le degré de satisfaction. Quant aux "*Fulfillment Theories*", elles assument que la satisfaction découle des effets réels quantifiables d'un épisode de soins de santé. Par opposition aux précédentes, ces théories de la satisfaction ne s'édifient pas sur la base d'un standard psychologique personnel pré-existant. Cependant, tout comme le modèle de Linder-Pelz (1982a, 1982b), ces deux types d'approches théoriques ne sont que faiblement supportées par les recherches empiriques.

D'après Pascoe (1983), cette situation s'explique par le fait que les concepts qui entrent en jeu dans ces divers modèles ne sont pas suffisamment bien circonscrits, notamment celui des attentes du client. Pour combler cette lacune, Pascoe (1983) de même que McMillan (1987) proposent d'emprunter les approches développées pour la mise en marché des biens et services commerciaux. Pour Pascoe (1983), la satisfaction se définit comme la réaction du client par rapport à des aspects saillants de l'environnement, du processus et

des résultats d'une expérience de soins. Le jugement du client est tributaire d'un standard personnel subjectif pouvant correspondre à une situation idéale, à la perception de ce qui est mérité, à ce qui est attendu d'après les expériences passées ou encore à un niveau jugé comme minimalement acceptable. Dans ce processus complexe de structuration de la satisfaction, des composantes affectives et cognitives sont impliquées. Tout comme les précédents, le modèle de Pascoe n'a pas été testé de sorte que sa validité reste à démontrer.

Pour terminer cette revue des modèles de satisfaction, il importe de souligner quelques autres travaux théoriques dans lesquels la satisfaction est abordée comme la variable médiatrice entre les soins (variable indépendante) et le comportement du client (principale variable dépendante) (Swan, Van Matre & McGee, 1985; Woodside, Frey & Daly, 1989). Dans cette optique, la satisfaction représente le dispositif intermédiaire par lequel les soins influencent le comportement subséquent du patient. Le pourcentage de variance expliquée par ce modèle (approximativement 40 %) confirme dans une certaine mesure la pertinence de l'approche. Cependant, les mécanismes précis qui contribuent à former le jugement du client et ses réactions ne sont pas élucidés pour autant.

Études empiriques

Depuis une dizaine d'années dans le domaine de la réadaptation, on compte un nombre accru de publications empiriques portant sur le thème de la satisfaction. Se situant dans une perspective évaluative, les recherches recensées ciblent, du plus général au plus spécifique, les programmes d'intervention (Daly & Flynn, 1985; Datta & Ariyaratnam, 1996; Davies & Hobbs, 1989; Elliott-Burke & Pothast, 1997a, 1997b; Heinemann et al., 1997; Koch & Merz, 1995; McComas,

Kosseim & Macintosh, 1995; Wilson, Crupi, Greene, Gaulin-Jones, Dehoux & Korol, 1995; Winter & Keith, 1988) et leurs impacts sur la qualité de vie (Fuhrer, Rintala, Hart, Clearman & Young, 1992; Hazard, Haugh, Green & Jones, 1994), les pratiques professionnelles (Chang & Hasselkus, 1997; Roush, 1995) ainsi que des modalités spécifiques de traitement telles que les aides techniques (Hachisuka, Dozono, Ogata, Ohmine, Shitama & Shinkoda, 1998; Korpela, Seppänen & Koivikko, 1993; Mann, Hurren & Tomita, 1993).

Dans la majorité des recherches examinées, la satisfaction est considérée comme une variable dépendante, reflétant les résultats d'une intervention donnée. Dans cette évaluation, d'autres indicateurs de résultats peuvent ou non s'adjoindre à la satisfaction, par exemple le statut fonctionnel (Fuhrer et al., 1992; Mann et al., 1993), la fréquence d'utilisation (Finlayson & Havixbeck, 1992; Kohn, LeBlanc & Mortola, 1994; Korpela et al., 1993) et la douleur (Hazard et al., 1994). En outre, la satisfaction peut être perçue comme une variable indépendante, à partir de laquelle on souhaite prédire le comportement du client (Elliott-Burke & Pothast, 1997b; Simon & Patrick, 1997). Selon les circonstances, le comportement attendu peut consister à se conformer au traitement recommandé, à faire une utilisation optimale des ressources fournies, à revenir consulter au besoin, ou encore à diffuser une image positive de la réadaptation dans l'entourage (Aharony & Strasser, 1993). En réadaptation, peu d'études se sont attardées à vérifier empiriquement l'existence d'un tel lien de causalité entre le degré de satisfaction et les comportements subséquents du client, bien que cette hypothèse s'avère, dans une certaine mesure, supportée par quelques résultats de recherche dans le domaine de la santé (Hall, Milburn & Epstein, 1993; Swan et al., 1985; Zastowny, Roghmann & Cafferata, 1989). Il

demeure que, pour les instigateurs des études réalisées en réadaptation, la perception d'un continuum entre l'intervention, le degré de satisfaction et un comportement souhaitable ressort comme un principe directeur implicite motivant la mise en application de la plupart des projets de recherche (Simon & Patrick, 1997).

Par rapport à la structure de la satisfaction, deux tendances peuvent être dégagées de la revue de la littérature empirique. Dans quelques études, la satisfaction est appréhendée de façon globale (Bynum & Rogers, 1987; Chamberlain & Gallop, 1988; Hazard et al., 1994; Mann et al., 1993). Dans ce contexte, le client est invité à donner une appréciation générale des services ou de l'intervention, ou encore à répondre à quelques questions construites pour cerner le phénomène dans son ensemble. Cependant dans la majorité des études (Bilodeau, 1994; Davies & Hobbs, 1989; Elliott-Burke & Pothast, 1997a; Finlayson & Havixbeck, 1992; Garber & Gregorio, 1989; Kohn, Enders, Preston & Motloch, 1983; McComas et al., 1995; Wilson et al., 1995), la satisfaction est évaluée par rapport à plusieurs aspects distincts des services ou de l'intervention. Les principales dimensions de la satisfaction spécifique à la réadaptation n'ont pas été précisément identifiées par des études. Cependant, Keith (1998) de même que Simon et Patrick (1997) considèrent qu'elles sont sensiblement les mêmes que dans les autres domaines de la santé. Ainsi, elles regroupent: 1) la relation client/personnel; 2) la qualité des soins; 3) la compétence technique des professionnels; 4) l'accessibilité des soins et des services; 5) la continuité de soins et des services; 6) les aspects financiers; et 7) l'environnement physique. En ce qui a trait aux aides techniques, des caractéristiques techniques plus précises, telles que le confort, la sécurité, la

facilité d'utilisation et d'autres critères d'évaluation ont été proposés par Batavia et Hammer (1990).

La prise en considération des méthodologies employées pour mesurer la satisfaction des clients de la réadaptation fait ressortir des lacunes concernant la construction des instruments de mesure et les échelles de cotation utilisées. Jusqu'à récemment, la majorité des études traitant de la satisfaction étaient basées sur des questionnaires construits spécifiquement pour les fins des recherches en cours (Hammel, 1996; Keith, 1998). Dans ces circonstances, peu de données permettaient d'apprécier leurs propriétés de mesure, les études de fidélité et de validité n'ayant pas été réalisées. De plus, les comparaisons entre les programmes d'aides techniques sont impossibles. Récemment, quelques chercheurs (Osborn, 1998; Simon & Patrick, 1997) ont tenté de contrer cette tendance à multiplier les échelles de satisfaction à usage unique. Leurs travaux de synthèse à orientation méthodologique permettent d'encadrer la démarche de ceux qui se lancent dans la construction d'un instrument de mesure de la satisfaction, afin de le rendre plus conforme aux règles de la métrologie.

Quant aux échelles de cotation, elles sont soit dichotomiques (Bynum & Rogers, 1987; Finlayson & Havixbeck, 1992; Mann et al., 1993), soit ordinales (Bilodeau, 1994; Datta & Ariyaratnam, 1996; Davies & Hobbs, 1989; Garber & Gregorio, 1990; Hazard et al., 1994; Wilson et al., 1995) incluant quatre, cinq et jusqu'à 10 niveaux de satisfaction. Dans la majorité des études recensées, le type d'échelle à réponse directe prévaut. Par exemple, les chercheurs déterminent des niveaux de réponses tels que: 1 = très satisfait, 2 = assez satisfait, 3 = plus ou moins satisfait, 4 = peu satisfait et 5 = pas satisfait du tout. Par ailleurs, on trouve des questionnaires utilisant un type d'échelle à réponses indirectes, où le

mot "satisfaction" n'apparaît pas, mais qui sont interprétées comme tel (Datta & Ariyaratnam, 1996; Kohn et al., 1994; Korpela et al., 1993; McGrath, Goodman, Cunningham, MacDonald, Nichols & Unruth, 1985). Dans le domaine de la réadaptation, l'impact des deux alternatives sur la fidélité, la validité et la sensibilité des mesures n'a pas été vérifié.

Quelque soit la méthode employée, les cotes de satisfaction moyennes s'avèrent très élevées. Par exemple, Wilson et ses collaborateurs (1995) rapportent que 97% des sujets évalués sont satisfaits d'une clinique ambulatoire de mobilité et de positionnement. En situation réelle, il est pourtant peu probable que tous ressentent et expriment un degré de satisfaction similaire à l'égard d'une panoplie de services. Pour que la démarche évaluative soit efficace, il est essentiel que les échelles de cotation soient suffisamment sensibles (Simon & Patrick, 1997).

Cependant, quelques études présentent des résultats de satisfaction plus nuancés (Bilodeau, 1994; Smedley, 1990). Par exemple, Bilodeau (1994) rapporte un degré moyen de satisfaction envers les prothèses pour le membre inférieur de 73.1 % avec un écart-type de 20. 3%. L'instrument de mesure de la satisfaction utilisé par Bilodeau, le "SAT PRO", a été éprouvé pour sa fidélité et sa validité (Bilodeau et al., 1998). De plus, l'attribution d'une prothèse constitue une intervention précise abordée, pour les besoins de l'évaluation, en regard de 14 dimensions distinctes. Ces facteurs semblent avoir favorisé l'émergence de résultats de satisfaction plus nuancés.

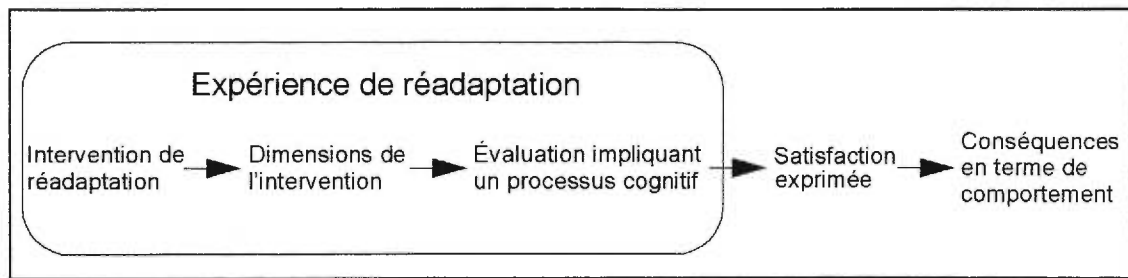
Conclusions

Ce survol des études empiriques et théoriques de la satisfaction démontre que l'un des concepts centraux de cette thèse manque de direction et d'ancrage théorique. Dans le domaine de la santé, quelques recherches de modélisation mettant en évidence les rôles prépondérants des attentes et des perceptions individuelles ont été menées. Toutefois, les hypothèses testées ne sont pas vérifiées et les résultats tendent au contraire à discréditer plusieurs des modèles proposés. Celui qui s'avère le plus prometteur aborde la satisfaction comme une variable intermédiaire, sans préciser toutefois les mécanismes spécifiques qui la sous-tendent. Par conséquent, peu de balises théoriques peuvent guider la recherche menée auprès des clients de la réadaptation et encore moins soutenir le développement d'un instrument de mesure apte à rendre compte de ce phénomène dans le champ des aides techniques.

Il reste qu'à partir des études empiriques, il est possible de dégager quelques principes généraux et une modélisation de variables respectant les idées maîtresses des travaux théoriques et empiriques révisés jusqu'à présent. Dans le contexte de la présente thèse, cet exercice permet de clarifier la compréhension du phénomène de satisfaction et de donner un cadre général au développement d'un instrument de mesure.

Tel que schématisé à la figure 1, les relations entre les variables liées à une expérience de réadaptation sont abordées et présentées de façon linéaire. La satisfaction est perçue comme une réaction découlant d'une intervention de réadaptation (variable dépendante) ainsi qu'en tant que stimulus à un comportement ou à une action consécutive du client (variable indépendante).

Figure 1: Modélisation de la satisfaction en réadaptation .



* Représentation de la satisfaction inspirée de Simon & Patrick (1997)

L'évaluation peut porter sur un programme, une modalité de traitement ou un service professionnel. Dans tous les cas, le phénomène étudié est décomposé en diverses dimensions correspondant à des aspects plus précis de l'intervention. Ce principe de multidimensionalité ressort clairement des travaux effectués dans le domaine de la réadaptation et se retrouve aussi dans les autres disciplines de la santé (Fitzpatrick, 1991, Hsieh & Kagle, 1991; McDaniel & Nash, 1990; Steiber & Krowinski, 1990).

Conformément aux travaux recensés, l'évaluation de ces dimensions implique un processus cognitif, dans lequel entre une part d'arbitraire. Ainsi, pour plusieurs chercheurs, la satisfaction se conçoit comme un processus dynamique, lié au style de vie, aux attentes, aux expériences personnelles, aux valeurs individuelles et aux valeurs sociales du client (Carr-Hill, 1992; Elbeck, 1987; Fitzpatrick & Hopkins, 1983; Hsieh & Kagle, 1991; Linder-Pelz, 1982a, 1982b; Strasser & Davies, 1991; Vuori, 1991). Dans cette optique, il pourrait s'avérer pertinent de permettre aux clients de définir les dimensions qui influencent leur satisfaction et de leurs accorder une pondération variable (Carr-Hill, 1992; Simon & Patrick, 1997).

En accord avec les principes généraux énoncés ci-haut, la satisfaction résultant d'une expérience de réadaptation se définit comme une évaluation positive ou négative de diverses dimensions d'un phénomène étudié. Cette évaluation est déterminée, dans une certaine mesure, par les attentes, les perceptions, les attitudes, les standards personnels et les processus de comparaison individuels.

3. Propriétés métrologiques requises des instruments de mesure

Fidélité

Les instruments de mesure sont généralement évalués en fonction de leur fidélité et de leur validité. La fidélité réfère à la capacité d'un instrument de mesure de reproduire un résultat de façon consistante lors de son utilisation en des occasions séparées. D'après Contandriopoulos, Champagne, Potvin, Denis et Boyle (1990), trois grandes approches permettent de cerner la fidélité d'un instrument, soit sa stabilité test-retest, sa reproductibilité interjuges et l'homogénéité de ses composantes.

D'abord, la stabilité des mesures peut être évaluée en comparant le résultat obtenu lors de l'utilisation de l'instrument à différents moments; il s'agit de la fidélité test-retest. La deuxième approche consiste à préciser dans quelle mesure les résultats obtenus par différents examinateurs avec le même instrument, auprès du même sujet et à un même moment sont reproductibles. On parle alors de fidélité inter-juges. Lorsque les données sont recueillies par entrevue, les juges ne peuvent questionner le même sujet simultanément. Par conséquent, les prises de mesures sont décalées et il est plus juste de parler de reproductibilité en deux temps (Feinstein, 1987). Les procédures découlant des approches précédentes peuvent s'appliquer à l'échelle de mesure dans son

ensemble, ou le cas échéant, pour chacun des items pris individuellement. Comme estimateurs de la fidélité, les statistiques les plus fréquemment utilisées sont les coefficients de corrélation, par exemple, le coefficient de corrélation intraclasse (Armstrong, 1981; Traub, 1994; Shrout & Fleiss, 1979), et les coefficients de concordance, telle que la statistique Kappa (Maclure & Willett, 1987). Dans les cas de données catégorielles et ordinales, Landis et Koch (1977) suggèrent d'interpréter les coefficients en fonction des catégories suivantes: pauvre (< 0.00), faible (0.00-0.20), acceptable (0.21-0.40), modéré (0.41-0.60), important (0.61-0.80) et presque parfait (0.81-1.00).

Finalement, lorsque la mesure se compose de plusieurs items, sa fidélité peut s'apprécier par le degré de consistance interne entre les items qui représentent diverses facettes d'un même concept. Pour estimer l'homogénéité d'un instrument de mesure, le coefficient alpha de Cronbach (Crocker & Algina, 1986) est couramment employé. D'après DeVellis (1991), les repères permettant d'interpréter les valeurs obtenues sont les suivants: inacceptable (< 0.60), indésirable (0.60-0.65), minimalement acceptable (0.65-0.70), respectable (0.70-0.80) et très bon (0.80-0.90). Lorsque les valeurs excèdent largement 0.90, DeVellis (1991) recommande de reconsidérer le nombre d'items dans le but de le réduire.

Validité

L'appréciation de la validité d'un instrument d'évaluation consiste à porter un jugement sur sa capacité à mesurer le concept étudié. Plus précisément, cette qualité réfère à l'adéquation qui existe entre le concept théorique et la façon de le rendre opérationnel par les variables retenues (Contandriopoulos et al.,

1990). La validité est donc fortement dépendante du cadre théorique utilisé et de l'état des connaissances dans un domaine d'étude. De plus, elle n'est ni immuable ni absolue mais dépend étroitement du contexte d'utilisation de l'instrument (Contandriopoulos et al., 1990; Silverstein et al., 1992). Par exemple, le PIADS a été validé auprès de personnes anglophones utilisant des verres correcteurs et des lentilles cornéennes (Day & Jutai, 1996). Il est peu probable que cet instrument puisse s'appliquer à une population francophone d'utilisateurs en fauteuil roulant, malgré les preuves de validité accumulées jusqu'à présent.

Ainsi, la qualité des interprétations que l'on peut inférer à partir des résultats d'un instrument de mesure découle du processus de validation utilisé (Messick, 1980). Le plus souvent, une seule étude de validation ne suffit pas. Au contraire, les résultats obtenus à l'aide de plusieurs devis de recherche doivent s'accumuler et supporter de manière cohérente le réseau de relations théoriques sous-jacent au concept évalué. Il s'ensuit que la validation d'un instrument de mesure est un processus perfectible et qui s'affirme à l'usage (Feldman, Haley & Coryell, 1990; Messick, 1980).

Classiquement, on distingue trois approches de validation des instruments de mesure, soit la validité de contenu, la validité de critère et la validité de construit (Contandriopoulos et al., 1990). La première, la validité de contenu, consiste à juger d'une part de la pertinence des éléments sélectionnés et, d'autre part, de leur capacité à représenter toutes les facettes importantes du concept à mesurer (Messick, 1980). Elle porte sur l'instrument ainsi que sur les items et implique, par conséquent, un développement minutieux des questions qui le composent. Une des procédures couramment employée pour estimer la validité du contenu

consiste à utiliser le jugement d'experts (Thorn & Deitz, 1989). Une deuxième approche de validation repose sur la corrélation entre les résultats produits par un instrument de mesure et un critère d'intérêt précis, tel qu'un comportement, un symptôme ou une mesure similaire. Des liens forts supportent l'utilité de cet instrument. Lorsque le critère se situe dans l'avenir, il est question de validité prédictive et lorsqu'il est contemporain, il s'agit de validité concurrente ou concomitante (Contandriopoulos et al., 1990; Messick, 1980).

La troisième approche de validation, appelée la validité de construit, s'intéresse directement aux relations entre les concepts théoriques et leur opérationnalisation (Contandriopoulos et al., 1990; Gompertz, Pound & Ebrahim, 1994). Elle comporte plusieurs formes et se réalise par divers types de stratégies. La première forme s'intitule la validité de trait et analyse la correspondance entre diverses façons de mesurer le même construit (Contandriopoulos et al., 1990; Messick, 1980). Le jugement qui s'ensuit concerne les deux instruments utilisés de sorte que la validation est symétrique. La deuxième forme, appelée validité nomologique, interpelle le réseau de relations théoriques entre le concept mesuré et d'autres concepts qui gravitent autour de celui-ci. Diverses hypothèses sur les résultats recherchés en administrant l'instrument sont proposées et testées. Par exemple, Bilodeau et ses collaborateurs (1998) administrent le SAT PRO en même temps qu'une échelle de dépression gériatrique dans le but de démontrer la relation entre les concepts de dépression et de satisfaction. Si elle est vérifiée, cette hypothèse ne prouve pas la validité de construit mais elle soutient la présomption que l'instrument mesure bien le concept étudié. Une troisième forme de validation de construit est la validation factorielle. Sa stratégie repose sur une procédure statistique qui permet de

regrouper les items qui composent l'instrument de mesure en un certain nombre de dimensions communes. Les énoncés qui relèvent théoriquement d'une même dimension conceptuelle devraient partager une proportion importante de variance commune (Contandriopoulos et al., 1990; Pedhazur & Schmelkin, 1991). D'après Harvey, Silverstein, Venzon, Kilgore, Fisher, Streiner & Harley (1992), le caractère unidimensionnel d'une mesure constitue une condition essentielle à sa validité.

Finalement, Ferketich (1991) considère qu'en plus d'être fidèle et valide, un instrument de mesure devrait être suffisamment sensible pour mettre en évidence des changements survenus entre des mesures successives d'un phénomène étudié. Il s'agit d'une propriété de mesure que l'on peut apprécier pour l'ensemble de l'échelle ou pour chacune de ses composantes. La sensibilité d'un instrument s'apprécie en fonction des objectifs pour lesquels il est utilisé. En effet, l'ordre de grandeur des changements considérés comme étant significatifs varie selon la nature des phénomènes mesurés de même qu'en fonction des interventions posées. Dans cette perspective, il est logique d'associer l'évaluation de la sensibilité à la démarche de validation de tout instrument de mesure.

4. Travaux préliminaires

Développement de l'ÉSAT

Les études présentées dans cette thèse reposent en partie sur des travaux antérieurs réalisés dans le cadre de la maîtrise de la candidate (Demers, 1995). Une revue de la littérature portant sur les aides techniques avait démontré qu'il n'existait pas d'instrument de mesure de la satisfaction des usagers, ni en

Amérique du Nord, ni en Europe. Dans ce contexte, une première version de l'ÉSAT avait été produite (Demers et al., 1996) et les paragraphes qui suivent en résument les étapes de construction.

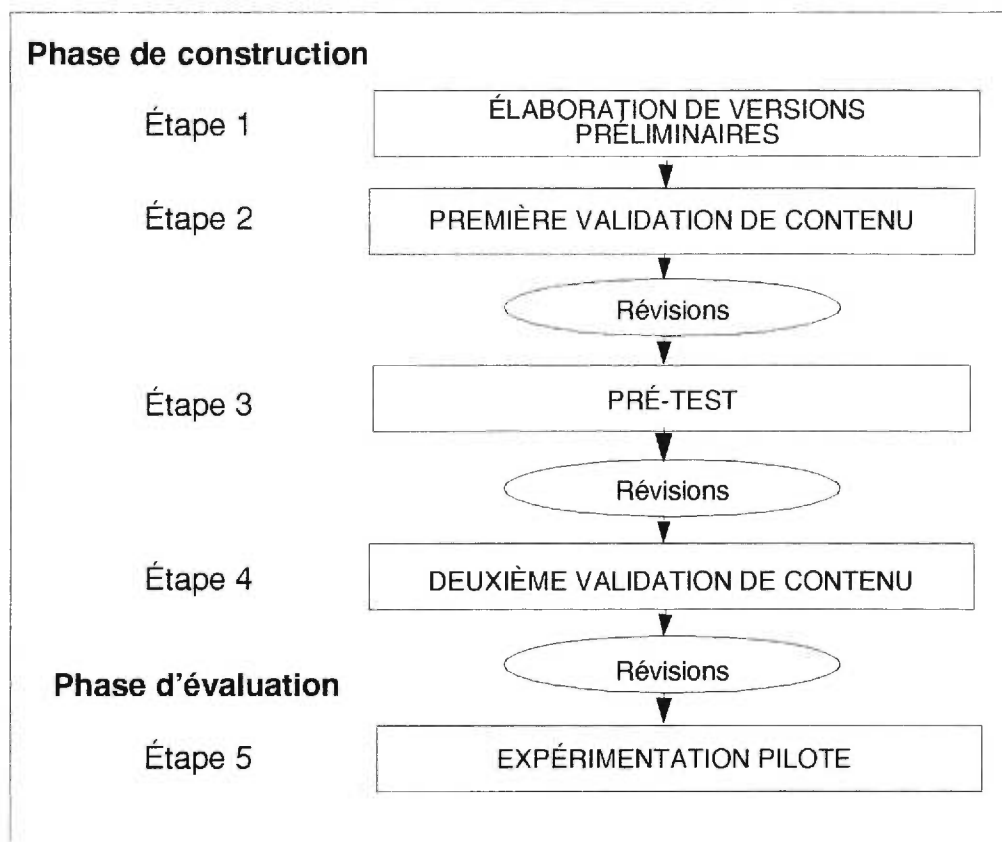
Pour l'élaboration d'un instrument de mesure, une démarche logique, rigoureuse et systématique est requise (Contandriopoulos et al., 1990). Dans le cas de l'ÉSAT, cette démarche avait été décomposée en quatre principales phases, conformes à celles décrites par Benson et Clark (1982). Les deux premières, correspondant à la planification et à la construction de l'instrument, avaient été complétées. La troisième phase, qui réfère à l'évaluation quantitative de l'instrument, n'avait été qu'amorcée avec la tenue d'une expérimentation pilote. La quatrième phase, la validation, a été réalisée dans le cadre de cette thèse.

La figure 2 représente les étapes reliées au développement de la version originale de l'ÉSAT. Celle-ci inclut: (1) l'élaboration de versions préliminaires, (2) la première validation de contenu par un groupe de 11 experts regroupant des chercheurs, des cliniciens, des administrateurs ainsi que des usagers, (3) le pré-test, (4) la deuxième validation de contenu et (5) l'expérimentation pilote auprès d'un échantillon de 30 sujets âgés utilisant divers types d'aides techniques reliées aux activités de la vie quotidienne. Chacune de ces étapes a été suivie par des révisions de l'instrument de mesure.

Description de l'ÉSAT

Le but de l'ÉSAT consiste à évaluer le degré de satisfaction d'une personne en regard d'une aide technique. Plus précisément, les objectifs sont: 1) de cerner le contexte dans lequel s'est développée la satisfaction ou l'insatisfaction de l'utilisateur envers son aide technique, 2) d'identifier les critères de satisfaction qui

Figure 2: Étapes de développement de l'ÉSAT*



* Reproduit du mémoire de maîtrise, 1995, p.29.

sont jugés comme étant importants pour un usager et 3) de connaître le degré de satisfaction à l'égard de chacun des critères de satisfaction ainsi que de façon globale.

La version expérimentale de l'ÉSAT qui a résulté des étapes de construction présentées ci-haut comprend 24 critères susceptibles d'influencer la satisfaction de l'utilisateur d'aides techniques. Ces critères appartiennent à trois catégories: la personne, l'environnement et la technologie. Par exemple, le confort, la durabilité et le poids relèvent de la catégorie technologie. Le service professionnel, l'entretien et la compatibilité technique se regroupent sous la catégorie environnement. Finalement, l'entraînement, l'apparence et l'effort

constituent des facteurs reliés à la catégorie personne. En somme, la satisfaction ou l'insatisfaction s'explique par de multiples sources d'influence, lesquelles agissent en interaction les unes avec les autres.

L'ÉSAT se compose de trois parties (voir appendice A). La première consiste à recueillir des renseignements généraux permettant de cerner le contexte d'utilisation de l'aide technique. La deuxième partie a pour but de déterminer l'importance attribuée à chacun des critères de satisfaction. En pratique, la tâche de la personne évaluée consiste à classer les fiches sur lesquelles s'inscrivent les noms des critères de façon à ce qu'elle identifie l'importance de chacun d'entre eux. À cette fin, un guide de classement comportant une échelle de type ordinal à cinq niveaux est utilisé: 1=pas important du tout, 2=peu important, 3=plus ou moins important, 4=assez important et 5 = très important. La troisième partie vise d'une part à préciser le degré de satisfaction envers les critères et, d'autre part, à explorer les sources d'insatisfaction. Pour cela, l'évaluateur demande à l'utilisateur de préciser son degré de satisfaction envers 20 des 24 critères précédents de même que de façon globale. Quatre items ont été retirés parce qu'ils étaient considérés comme ambigus et difficiles à coter avec l'échelle de satisfaction. Celle-ci est matérialisée sous la forme d'un boîtier rectangulaire comportant les niveaux suivants: 1=pas satisfait du tout, 2=peu satisfait, 3=plus ou moins satisfait, 4=assez satisfait et 5=très satisfait. L'utilisateur déplace le curseur du boîtier à la position qui correspond le mieux à son degré de satisfaction. Pour terminer, l'évaluateur interroge l'utilisateur sur tous les critères dénotant une satisfaction faible, c'est-à-dire ceux générant des scores de 1, 2 ou 3 sur l'échelle de satisfaction. Cette procédure permet de mieux comprendre les facteurs qui influencent la perception de l'utilisateur.

Les tâches et les questions constituant l'ÉSAT sont administrées dans le cadre d'une entrevue structurée qui dure approximativement 45 minutes. Ce mode d'évaluation permet de recueillir l'opinion des personnes pour lesquelles l'utilisation d'un questionnaire écrit est impossible en raison d'incapacités physiques.

5. Position du problème et objectifs généraux

Les études portant sur l'évaluation des effets et conséquences d'aides techniques ont le mérite de s'être penchées sur un domaine encore très peu étudié. Curieusement, malgré l'état morcelé de l'avancement des connaissances, les conclusions convergent pour soutenir le développement d'instruments de mesure reflétant le point de vue de l'utilisateur et notamment son degré de satisfaction. En partie pour des raisons conceptuelles, la mesure de ce phénomène pose un grand défi aux chercheurs. En effet, l'analyse des recherches théoriques indique qu'il s'agit d'un concept difficile à définir et dont les déterminants et mécanismes sous-jacents sont peu connus. Le fait que les fondements théoriques de la satisfaction soient mal circonscrits n'empêche pas la prolifération d'études empiriques dans le champ de la réadaptation. Cependant dans bien des cas, les méthodologies employées sont critiquables et freinent le développement des connaissances. En l'occurrence, aucun instrument fiable n'est disponible à l'heure actuelle pour cerner la satisfaction envers divers types d'aides techniques. C'est une contribution au développement d'un tel instrument que la présente thèse se propose de faire.

Pour y arriver, trois étapes principales sont prévues. Chacune de ces étapes est rapportée, sous forme d'article, dans les chapitres qui suivent. Dans un premier

temps, les résultats d'une étude internationale portant sur la validité de contenu de l'ÉSAT sont présentés. Les travaux préliminaires ayant été réalisés dans le contexte d'un seul milieu culturel, soit celui de Montréal, des preuves additionnelles de validité de contenu sont requises pour vérifier si l'instrument convient à d'autres types d'environnements. Après avoir clairement établi les forces et limites du contenu de l'ÉSAT, la démarche méthodologique s'est poursuivie en évaluant sa fidélité. Cette propriété de mesure a été appréhendée pour chacun des énoncés pris individuellement et ce, pour deux approches de fidélité, soit la stabilité test-retest et la reproductibilité interjuges. Le deuxième article relate cette importante étape de la démarche. Finalement, les résultats de ces étapes permettent d'approfondir l'analyse des items et de vérifier si un sous-ensemble d'énoncés présente des caractéristiques optimales de fidélité et de validité. Cette étude fait l'objet du troisième article. Celui-ci comporte deux séries d'analyses. La première série évalue le rendement des items en regard de six critères de mesure alors que la deuxième série raffine cette évaluation en soumettant les items les plus performants à des analyses factorielles. La question de la structure conceptuelle de la satisfaction envers les aides techniques y est également abordée. Après ces trois articles, qui forment le coeur de la thèse, la réflexion amorcée sera approfondie dans le cadre d'une discussion générale qui permettra au lecteur de replacer ces trois articles dans un contexte élargi.

Cette thèse contribue à l'avancement des connaissances de plusieurs façons. Principalement, elle permet de produire une version améliorée et testée d'un instrument de mesure des résultats des aides techniques, ce qui répond à une demande pressante en Amérique du nord et en Europe. Ensuite, elle offre un

modèle méthodologique détaillé de développement d'une mesure de la satisfaction, ce qui est utile pour le domaine de la réadaptation. Finalement, la question de la conceptualisation de la satisfaction envers les aides techniques est abordée ici pour la première fois.

Chapitre 2

An international content validation of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST)

**An international content validation of the Quebec
User Evaluation of Satisfaction with assistive
Technology (QUEST)**

Louise Demers, M.Sc., OT(C)*

École de réadaptation, Université de Montréal, Canada et
Centre de recherche de l'Institut universitaire de
gériatrie de Montréal, Canada

Roelof D. Wessels, B.Sc.

Institute for Rehabilitation Research, IRV, Hoensbroek, The Netherlands

Rhoda Weiss-Lambrou, M.Sc., OT(C)

École de réadaptation, Université de Montréal, Canada

Bernadette Ska, Ph.D

Centre de recherche de l'Institut universitaire de
gériatrie de Montréal, Canada

Luc P. de Witte, Ph.D

Institute for Rehabilitation Research, IRV, Hoensbroek, The Netherlands

Article to be published in:

Occupational Therapy International, 6(3), 1999.

* Corresponding author

Louise Demers

Centre de recherche de l'Institut universitaire de gériatrie de Montréal

4565, Chemin Queen Mary

Montreal (Quebec), Canada, H3W 1W5



Abstract

The Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) is an outcome assessment tool designed to identify the sources of user satisfaction and dissatisfaction with assistive technology. The purpose of this paper is to present the results of an international content validation study of QUEST that was conducted as part of the ongoing development of this assessment instrument. A questionnaire developed to measure the relevance of the QUEST items and to critique the administration procedures was completed by an international group of 12 assistive technology experts from the United States, The Netherlands and Canada. At the test level, the findings revealed that QUEST was adequately sampled in terms of embracing all the important facets of satisfaction with assistive technology. At the item level however, it was revealed that changes needed to be made to the wording of the satisfaction variables and to the procedures used in administering QUEST in order to ensure optimal content validity. A limitation of the study is the heterogeneity of the testing sites that may have reduced the chances of reaching consensual agreement among the content experts. It is recommended that future studies consider alternative methods such as factorial analysis to support the findings of this study.

Keywords: Assistive technology; Content validity; User satisfaction; International collaboration; Outcomes assessment; Instrument development.

Introduction

In the current context of quality assurance, cost-containment and consumerism, the need for meaningful and reliable outcome measures of assistive devices has become of foremost importance to assistive technology providers in North America (DeRuyter, 1995; Scherer & Vitaliti, 1997; Trachtman, 1994) and in Europe (Andrich, Ferrario, & Moi, 1997; Lorentsen, 1997). There is a need to evaluate the effectiveness, efficiency and consumer satisfaction of assistive devices that are widely used by individuals with disabilities. As a predictor of use, consumer satisfaction is hypothesized to have a strong impact on cost issues that are significant for long term provision of assistive technology (Zastowny, Roghmann & Cafferata, 1989). Satisfaction is also closely linked with the concept of quality of life, which can be strongly affected by the disablement process (Canadian Burden of Illness Study Group, 1998; Verbrugge & Jette, 1994; Vuori, 1991). There is a general agreement today among assistive technology professionals on the need and importance of measuring user satisfaction with assistive devices. Valid results however, depend upon appropriate instruments and methodologies (Smith, 1996).

Based on the fact that no such instrument was available, The Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) (Demers, Weiss-Lambrou & Ska, 1996) was developed as a measure of user satisfaction with a wide range of assistive devices. This experimental version of QUEST was designed to identify the sources of user satisfaction and dissatisfaction with assistive technology and was developed both for clinical and research purposes. In the clinical setting, QUEST may serve as a tool to evaluate the outcomes of service delivery by identifying those important characteristics of the user's

device, the service delivery of the technology or the person's living environment that are problematic and unsatisfying. In a research context, the instrument may serve to monitor the quality of service delivery systems and procedures using results gathered from various categories of assistive device user groups. This test version of QUEST has not been subjected to all the controls required to appraise its measurement properties and justify its use. One important step in evaluating the adequacy of a test is to provide evidence of content validity. Several approaches can be appropriate. The considerable interest generated by QUEST in several countries prompted the use of the judgments of content experts. The purpose of this paper is to present the results of an international content validation study of QUEST that was conducted as part of the ongoing development of this assessment instrument.

Description of QUEST

The concept of satisfaction as defined in QUEST refers to a person's critical evaluation of those distinct dimensions of the assistive device that are influenced by his/her expectations, perceptions, attitudes and personal values. Its focus is on how satisfied the user is with certain characteristics of the assistive device. The experimental version of QUEST used in this study consisted of 24 distinct items and definitions (See Table I) related to the technical characteristics of the device as well as to several psychosocial and service delivery attributes. Each of the 24 QUEST items are factors that can possibly influence user satisfaction.

QUEST is divided into three parts and is administered in a face-to-face interview. The evaluator administers the assessment and records the client's answers on

Table I: QUEST items and definitions

Item	Definition
1. Simplicity of use	Ease in using the ATD.
2. Repairs/servicing	Ease in having the ATD repaired and serviced.
3. Maintenance	Simplicity of upkeep and care of the ATD by oneself.
4. Installation	Simplicity to assemble and/or set up the ATD.
5. Effectiveness	Goal achievement with the ATD.
6. Cost	Expenses associated with purchasing, maintaining and repairing the ATD.
7. Professional service	Quality of information on ATD provided, accessibility and competence of professionals.
8. Durability	Robustness and sturdiness of the ATD.
9. Multi-purposefulness	Possibility to adapt and use the ATD for multiple activities/purposes.
10. Adjustments	Simplicity in setting/fixing the components of ATD.
11. Comfort	Physical and psychological well-being associated with use of ATD
12. Service delivery	Ease in acquiring the ATD including length of time.
13. Follow-up services	Ongoing support services for ATD.
14. Appearance	Design, form, color and acceptability of the ATD.
15. Transportation	Convenience of transporting the ATD via the desired means of transportation.
16. Device compatibility	Suitability of the ATD with other used technologies.
17. Weight	Ease in lifting and/or moving the ATD.
18. Safety	Degree to which the ATD is safe, secure and harmless.
19. Dimensions	Convenience of the device's size (height, width, length).
20. Motivation	Incentive to use the ATD at home and in public.
21. Training	Learning how to use the ATD.
22. Social circle support	Support from family, peers & employer in using the ATD, whether physical or psychological.
23. Reaction of others	Positive and encouraging attitude of others.
24. Effort	Little physical or psychological exertion required in using the ATD

Legend: ATD Assistive Technology Device

the QUEST scoring sheet. Part one consists of 18 close-ended questions aimed at describing the context in which user satisfaction or dissatisfaction developed. In Part two, the evaluator asks the user to rate the degree of importance he or she attributes to those items considered applicable to the assistive device being evaluated. As QUEST was designed to assess a wide range of assistive technology devices, not all items are applicable to every user and every situation. Importance is rated on a 5-point scale, with a score of 1 denoting "of no importance" and 5 indicating "very important". The user is also invited to add any other satisfaction items considered important. In Part three of QUEST, the client is asked to rate his or her satisfaction with the same items. Each item is scored with a 5-point satisfaction scale, with a score of 1 denoting "not satisfied at all" and 5 indicating "very satisfied". For each item that is scored 3 ("more or less satisfied"), 2 ("not very satisfied") or 1 ("not satisfied at all"), the user is asked to comment or explain the source(s) of dissatisfaction. Finally, the client is asked to score his or her overall satisfaction with the assistive technology device.

This test version of QUEST was designed in a "card-playing" format version only and the assessment materials were presented as a kit. These materials consisted of a set of 24 playing cards (with each card representing a satisfaction item); an interactive assessment box for classifying the cards (with the 5-point scale of degree of importance on one side of the box) and rating the satisfaction items by moving a sliding bar (with the 5-point scale of degree of satisfaction on the reverse side of the box); an instruction manual and the QUEST scoring sheet. The instrument was developed simultaneously in French and English and a Dutch translation was later created (Wessels, De Witte, Weiss-Lambrou, Demers & Wijlhuizen, 1998).

The process of translation demands great attention and a considerable investment of time. The production of the Dutch version of QUEST involved several steps. First, the authors of QUEST and the researchers from IRV (Institute for Rehabilitation Research) at Hoensbroek in the Netherlands discussed the various issues related to the development and application of the assessment tool. This dialogue of viewpoints began with a meeting at IRV, continued through electronic mail correspondence and subsequently led to certain changes to the original English and French versions of QUEST. Following this, the English QUEST was translated into Dutch and became known as D-QUEST. There was then a back-translation of D-QUEST into English by a professional translator who had a background in occupational therapy. The differences between the original and the back-translated English versions were discussed with the translator as recommended in the literature (Guillemin, Bombardier & Beaton, 1993). Finally, some minor changes to D-QUEST were made in cooperation with several Dutch occupational therapists. This last step was essential to adapting the instrument to the specific context of assistive technology provision and use in The Netherlands.

QUEST as an assessment tool for occupational therapists

Technological advances create new opportunities for persons with disabilities. In this context, occupational therapists are repeatedly faced with the difficult task of having to choose among a wide range of assistive technology devices (Bain, 1997; Post, 1993). Technology-based solutions are indeed available for problems related to the three domains of human occupation central to the

profession's practice that is, self-care, productivity and leisure (Canadian Association of Occupational Therapists, 1997).

Although user satisfaction with assistive technology is a timely and important subject of study, there is no standardized assessment tool currently available for this purpose. In some studies on the use of assistive technology, global measures of satisfaction have been obtained (Finlayson & Havixbeck, 1992; Mann, Hurren, & Tomita, 1993). However, results based on overall satisfaction ratings do not provide specific indications as to the sources of user dissatisfaction. For effective follow-up and improved service delivery, it is of utmost importance to use an assessment tool that encompasses the entire range of influences on a person's decision to use or not to use an assistive device. Occupational therapy clients present different degrees of disability and impairment, which challenge the use of paper and pencil assessment formats. Although QUEST is a standardized evaluation tool, its administration procedures can be adapted to meet the needs and abilities of the user with no severe cognitive disabilities; the assessment materials allow for flexibility in the way in which the user scores the items (i.e. by pointing or by responding verbally) while at the same time the tool is administered in a very structured manner so as to obtain reliable results that can be used for comparison purposes. The assessment is designed to be administered by practitioners working in the service delivery of assistive technology. When used for research purposes, therapists should not interview their own clients but rather those of colleagues so as to reduce researcher bias. Similarly, in a clinical context, their own therapist should not evaluate clients.

The experimental QUEST was developed and pre-tested in a single cultural setting, that of Montreal. The basic list of variables was based on literature on this subject (Batavia & Hammer, 1990; Phillips & Zhao, 1993; Ward, 1990) and clinical experience of authors. It was also discussed by a panel of experts during the construction phase of the instrument development to insure inclusion of all relevant aspects (Demers et al., 1996). Additional content validity evidence was needed because several changes to the assessment tool were made following the initial pre- testing. Moreover, the present content validation study was necessary to examine if the instrument was suitable for cross-cultural applications.

Methodology

Appraisal of an assessment instrument's validity consists of evaluating its capacity to measure the concept under study (Contandriopoulos, Champagne, Denis & Pineault, 1993; Harvey et al., 1992). It refers to the adequacy between the theoretical concept and its operationalization at the variable level. Classically, test developers are concerned with three types of validity: content, criterion-related and construct validity (Contandriopoulos et al., 1993). Content validity, the subject matter of this paper, refers to the extent to which the instrument covers the scope of the construct of satisfaction. It includes evaluating the relevance of selected components (content relevance) and their capacity to represent every facet of the measured concept (content coverage) (Messick, 1980). According to Guion (1977), five conditions are necessary for an instrument to be considered as having a valid operational definition on the basis of content validity: (1) the content needs to have a generally accepted meaning; (2) the domain must be unambiguously defined; (3) the content must be relevant

to the purposes of measurement; (4) the domain must be adequately sampled, and (5) the response content must be reliably observed and evaluated. In the present study, the opinions and expertise of assistive technology professionals referred here to as "content experts" were used to evaluate the degree to which QUEST measures the intended content area in terms of the first four of these conditions. The issue of QUEST's reliability properties will be investigated in a future study.

A volunteer sampling technique was used to recruit the content experts. In light of some of the publications on QUEST, a number of researchers and clinicians around the world contacted the authors in Montreal to inquire into the availability of the assessment tool. Thirteen of these persons were provided with QUEST kits and were given 6-8 months to use the satisfaction tool with any number of assistive devices available to them. They agreed to share their data with the Montreal authors and to evaluate the instrument's content, format and usefulness by completing a questionnaire. These content experts represented 11 facilities (university, clinical and industrial settings) in Canada, the United States and The Netherlands (see Table II). The French version of experimental QUEST was evaluated by experts in Montreal and Quebec City; the English version was evaluated in Toronto, Peaback (New Jersey), Pittsburgh, and Milwaukee; and the Dutch version was used at Hoenbroek and Leiden.

Based on similar work conducted by Reid and Jutai (1997), the questionnaire developed for this purpose consisted of 40 questions dealing with the following five issues: (1) the content coverage and item relevance, (2) the administration procedures, (3) the instruction manual, (4) the assessment materials and (5) the commercial publication of QUEST. Only the first two of these issues address the

instrument's content validity and are presented in this paper. The complete questionnaire however can be obtained upon request from the authors.

To evaluate the content coverage and item relevance, the content experts were first asked to rank the importance of each of the 24 items using a three-point scale: ("of primary importance", "of secondary importance" and "neither") and to comment on their answers. The option "neither" was to be selected if the expert felt that the item should not be included among the satisfaction variables. Following this ranking of the item relevance, the experts were asked three questions:


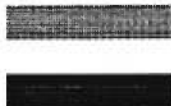

- Are there any other variables that you would like to see included in QUEST?
- Are there some variables that you consider not clear and/or difficult to explain?
- Do you find the definitions of the satisfaction variables useful?

For the purpose of appraising the administration procedures, the content experts were asked 10 questions related to the usefulness and comprehensibility of the QUEST tasks, the sequence of administering QUEST and the interactivity feature of the assessment tool.

Results

Although 13 content experts completed and returned the questionnaires, the responses of one expert were not included in the results of this study as QUEST had not been administered according to the procedures outlined in the manual. Of the 12 content experts who used and tested QUEST in the standardized manner, there were six university educators, four practitioners, one clinical

Table II: International content experts of QUEST

Facility	City	Country	Assistive Devices
<i>Institut de réadaptation de Montréal</i>	Montreal		Lower limb prostheses (n= 17) Powered wheelchairs (n=10) Manual wheelchairs (n=7)
<i>Centre de réadaptation Lucie Bruneau</i>			Powered wheelchairs (n=53) Manual wheelchairs (n=34)
<i>Centre de réadaptation Constance Lethbridge</i>			Powered wheelchairs (n=4) Manual wheelchairs (n=3)
Orthofab Inc.			Powered wheelchairs (n=15) Manual wheelchairs (n=5)
<i>Université Laval</i>	Quebec City	Canada	Environmental control devices (n=4)
Bloorview McMillan Center	Toronto		Environmental control devices (n=4)
Institute for Rehabilitation Research (iRv)	Hoensbroek	 The Netherlands	Toilet adaptations (n=100) Manual wheelchairs (67) Shower seats and chairs (n=66) Scooters (52) Powered wheelchairs (n=9) Adapted bed (n=28) Wide doors (n=24) Stairlift (n=13) Adjustable kitchens (n=11)
TNO Prevention and Health	Leiden		
Mathenay School & Hospitals (two experts)	Peaback (NJ)	 United States	Augmentative communication devices (n=8) Powered wheelchairs (n=15)
University of Pittsburgh	Pittsburgh (PA)		Manual wheelchairs (n=23)
University of Wisconsin	Milwaukee (WI)		Augmentative communication devices (n=6)

research director and one assistive technology program coordinator. Table II presents the facilities and cities where QUEST was used as well as the type and number of assistive devices evaluated. Since QUEST was designed as a generic instrument, its applicability for clinical and research purposes needed to be tested with more than one type of assistive technology device. QUEST was tested on a total of 578 assistive devices, which were primarily seating and mobility aids (wheelchairs, scooters) together with transfer aids (toilet adaptations, shower seats and chairs, stair-lifts), augmentative communication devices, lower limb prostheses and environmental control devices.

QUEST content coverage and item relevance

The relevance of the items to the domain of satisfaction was firstly assessed by asking the content experts to rate the importance of each of the 24 items using a three-point scale ("of primary importance", "of secondary importance" and "neither") and to comment on their answers. The experts rated all items except for one respondent who omitted six items that were judged to be redundant. To quantify the subjective judgments obtained, percentage of agreement was computed. Although this index does not account for the contribution of chance agreement, it is widely used, easy to compute, and relatively simple to interpret (Thorn & Dietz, 1989). As shown in Table III, the content experts expressed a high degree of agreement (>85%) on the primary importance of the following six items: "Simplicity of use", "Repairs/servicing", "Effectiveness", "Durability", "Comfort" and "Safety". A moderate degree of agreement (between 60% and 75%) was reached on the primary importance of five other items: "Professional services", "Weight", "Dimensions", "Training" and "Effort". A moderate degree of agreement (between 60% and 75%) was obtained on weak ratings of

Table III: Percentage of agreement on the relative importance of items

	Item	Primary importance (%)	Secondary Importance (%)	Neither (%)
High agreement on primary importance of items	1. Simplicity of use	12 (92%)	1 (8%)	0
	2. Repairs/servicing	11 (85%)	2 (15%)	0
	5. Effectiveness	12 (92%)	1 (8%)	0
	8. Durability	11 (85%)	2 (15%)	0
	11. Comfort	11 (85%)	2 (15%)	0
	18. Safety	12 (92%)	1 (8%)	0
Moderate agreement on primary importance	7. Professional service	8 (67%)	4 (33%)	0
	17. Weight	8 (62%)	5 (38%)	0
	19. Dimensions	8 (67%)	4 (33%)	0
	21. Training	9 (69%)	4 (31%)	0
	24. Effort	9 (75%)	3 (25%)	0
Moderate agreement on weak ratings of importance	4. Installation	5 (38%)	3 (23%)	5 (38%)
	23. Reaction of others	4 (31%)	7 (54%)	2 (15%)
Weak agreement for all given options	3. Maintenance	7 (54%)	6 (46%)	0
	6. Cost	6 (46%)	6 (46%)	1 (8%)
	9. Multi-purposefulness	6 (46%)	5 (38%)	2 (15%)
	10. Adjustments	7 (54%)	5 (38%)	1 (8%)
	12. Service delivery	6 (50%)	5 (42%)	1 (8%)
	13. Follow-up services	6 (50%)	5 (42%)	1 (8%)
	14. Appearance	6 (46%)	7 (54%)	0
	15. Transportation	7 (54%)	6 (46%)	0
	16. Device compatibility	6 (50%)	5 (42%)	1 (8%)
	20. Motivation	7 (54%)	4 (31%)	2 (15%)
	22. Social circle support	6 (46%)	6 (46%)	1 (8%)

importance for two items, namely "Installation" and "Reaction of others". For the remaining 11 items, opinions were more divergent and agreement was weak for all the given options.

In total, 55 comments on the QUEST items were provided to explain the experts' ratings. These comments were classified into three categories: pertinence of the items, wording of the items and redundancies. A fair proportion of the comments (n=23) were related to the pertinence of the items comprising the test. Approximately half of these comments (n=13) revealed the reasons for the weak ratings of importance. For example, one expert wrote that satisfaction with "Service delivery" was non-applicable to that particular context of study; another expert felt that this item did not address the domain of assistive technology per se. Some comments (n=3) emphasized the primary importance accorded to certain items such as, "Repairs and servicing" which was considered to be a very important consumer criterion. The remaining comments (n=7) of this first category challenged the applicability of some items for specific devices, clients, and service delivery systems. For example, in one respondent's opinion, the item "Adjustments" was important but felt that the user should not perform the adjustments of the device.

The second category of comments addressed the problem of wording or phrasing of certain items. A fair number of the remarks (n=23) focused on this issue and were congruent with the answers to subsequent questions on wording clarity and general meaning of the items. In the last and smallest group of comments (n=9), the experts reported some overlapping and redundancy across items. For example, two experts felt that "Simplicity of use" and "Effort" tap into the same aspect of satisfaction with assistive devices. Similarly, "Professional

services", "Follow-up services" and "Repairs/servicing" were considered to be closely related. No agreement higher than 17% (n=2 judges) was however obtained on identifying which items overlapped.

The next question of the questionnaire sought to establish whether or not the domain of user satisfaction with assistive technology had been adequately sampled in terms of comprising all the important facets. The content experts were asked, "Are there any other variables that you would like to see included in QUEST?" Two response options were possible: "yes" and "no". If the answer was in the affirmative, the additional items were identified. Nine experts (75%) considered that no other items needed to be added to QUEST. Three experts offered suggestions of items that were device-specific.

For the purpose of examining if the satisfaction items had a generally accepted meaning, the content experts were asked, "Are there some variables that you consider not clear and/or difficult to explain?" As previously, a "yes" or "no" response choice was offered followed in the affirmative by precise item identification. Results revealed that several QUEST items were, according to collective opinion (100% "yes"), not clear or difficult to explain. The wording or phrasing of some items was considered to be sometimes awkward and much was left to the evaluator's interpretation and discretion to use simpler language or provide device-specific examples. Experts emphasized the need for using simple and clear wording of the items bearing in mind that device users are generally cognitively intact but have wide-ranging cultural and educational backgrounds. The experts rated certain items as being difficult-to-explain. The most frequently cited items were "Motivation" and "Social circle support" pointed out by six experts (50%), "Multipurposefulness", "Device compatibility" and

"Effort" identified by five experts (42%), "Effectiveness" selected by four experts (33%) as well as "Maintenance", "Installation" and "Professional services" reported by three experts (25%). The remaining 15 items were found to have a generally accepted meaning as they were never or seldom considered as being difficult to explain.

Finally, the last question of this part of the questionnaire inquired into the usefulness of the item definitions. Using a three-point scale ("very useful", "somewhat useful", "not useful at all") the content experts were asked to rate how useful they felt were the definitions and to comment on their answers. Seventy-five percent (n=9) of the respondents felt that the item definitions were very useful. One expert commented that the definitions could be further developed to apply to different types of assistive devices. The "Somewhat useful" option was chosen by two experts (17%) and only one expert considered the definitions were "Not useful at all".

Administration procedures

Evaluation of the administration procedures of QUEST was addressed through 10 questions related to the usefulness and comprehensibility of the QUEST tasks, the sequence of administering QUEST and the interactivity feature of the assessment tool.

The first group of questions related to the card sorting task used to assess the user's perceived importance of the satisfaction items. The convenience of the task was rated on a 3-point scale ("very useful", "somewhat useful", and "not useful at all"). Since one expert did not complete this section of the questionnaire, the results of 11 experts were analyzed. Eighty-two percent of the

experts (n=9) felt that this task was "somewhat useful" or "not useful at all". The reasons put forward were that some subjects were unable to manipulate the cards or that the task was considered unnecessary, repetitive and childish. For those subjects with severe disabilities, four content experts (36%) considered this task difficult. Seven (64%) qualified the task as easy without specifying severity of disability. A majority of experts (n=8, 73%) indicated that this task did arouse some interest and helped keep the client focused.

Similarly, the usefulness of the moving sliding bar that was used to indicate the degree of satisfaction with each of the items was appraised by the content experts with a 3-point scale ("very useful", "somewhat useful", and "not useful at all"). Seven experts (64%) found this task "as somewhat useful" or "not useful at all" while the remaining four (36%) considered it to be "very useful". Some experts explained that subjects generally preferred to rate their satisfaction by answering verbally or by pointing to their responses rather than by repeatedly moving the scale bar. While some subjects found the task childish and embarrassing, others enjoyed it. Eighty-two percent of the experts (n=9) felt that the sliding bar was easy to manipulate; only two experts (18%) reported that it was physically too demanding for some clients. As expected, no agreement was reached as to the appeal of the sliding bar task for the clients, with six experts in favor (55%) and five against (45%).

The other group of questions in this section evaluated some of the general features of QUEST, the sequence of administration and the interactivity feature of the assessment tool; all 12 experts completed this section of the questionnaire. The experts were asked if the tasks were clearly understood by the clients. Opinions on this question were split in half; the main problem

seemed to be that some subjects had difficulty distinguishing between "importance" and "satisfaction". The experts were also asked if they felt there should be a change in the sequence of administering QUEST. The majority (n=8, 73%) liked the sequence as it is while only three experts (27%) proposed minor modifications; one expert failed to answer this question. An example of a proposed sequence change was to have the client attribute the satisfaction score and identify the source of dissatisfaction at the same time before moving on to the next item. A good agreement among the experts was reached (n=10, 83%) on the usefulness of the interactive feature of QUEST; some content experts explained that the interactivity feature was useful because it allowed them to confirm that the items were understood by the clients. Finally, the need for other QUEST formats was evaluated. Self-report questionnaires, phone surveys and computer versions were suggested respectively by five (42%), four (33%) and one (8%) content expert.

Discussion

The use of content experts is a common and practical approach to content validation of occupational therapy assessments (Law, 1987; Thorn & Deitz, 1989). It was particularly relevant in the development of QUEST to seize the opportunity created by the international interest for this assessment tool and to build on the shared experiences and opinions of assistive technology experts of several countries. Participants in this international network were able to work in close collaboration with the QUEST authors. Through electronic mail, difficulties related to geographical distances between testing sites were considerably reduced and communication was easy, quick and accessible to all. In addition, international conferences such as those held by the World Federation of

Occupational Therapists (WFOT) in Montreal (1998), the Rehabilitation Engineering and Assistive Technology Society of North America (RESNA) in Pittsburgh (1997) and Minneapolis (1998), as well as the Telematics application for the Integration of the Disabled and the Elderly (TIDE) conference in Helsinki (1998) allowed several of the experts to meet with the QUEST authors as well as with other content experts.

Cross-national bias may exist in measurement (Guillemin et al., 1993). Differences in the meaning of language can lead away from the intended purpose of the items and rating instructions and therefore, when an instrument is standardized for use with different cultural groups, word-by-word translation is not sufficient (Fisher, Liu, Craig & Pan, 1992; Vallerand, 1989). In translating QUEST from English to Dutch (D-QUEST), several Dutch assistive technology provision counselors were involved. This procedure is strongly recommended in the literature when adapting an instrument to the specific cultural context where it will be used (Fisher et al., 1992; Vallerand, 1989). Moreover, the discussions between the Hoensbroek and the Montreal researchers during the translation process permitted covert meanings to be overtly challenged and debated. In order to compare results, it was of critical importance that the translated QUEST versions measured the exact same scope of the concept of satisfaction in French, English, and Dutch.

The nature of the interpretation inferred from gathered results ensues from the kind of validity evidence demonstrated (Messick, 1980). The results of the present study were obtained from more than one cultural setting. This distinctive feature of the study is an important advantage for the cross-cultural application of QUEST. On the other hand, the heterogeneity of the testing sites in terms of

their type of culture, assistive devices and clients may have reduced the chances of reaching consensual agreement among the content experts. As the opinions expressed did not always support each other, the results should be interpreted carefully before making important changes to QUEST. Notwithstanding this limitation of the study, some very meaningful results were obtained.

One of the necessary conditions for content validation of a measuring instrument is the appropriate sampling of its content. The results of this study revealed that QUEST has been adequately sampled in terms of comprising all the important facets of satisfaction. In fact, very few additional items were proposed. Approximately half of the items comprising QUEST were strongly supported by the experts as valid dimensions of satisfaction with assistive technology. On the other hand, some items were surprisingly not considered to be of primary importance by the experts and this perhaps was because the items were less generic but yet very important for certain types of assistive devices. For example, the item "Transportation", defined as convenience of transporting the assistive device via the desired means of transportation, did not really apply with toilet adaptations, shower seats, or environmental controls but yet this satisfaction variable might be very important for wheelchairs users. Only two items were ranked poorly as regard to importance by an important number of content experts, namely "Installation" and "Reaction of others". The item "Installation" was defined as simplicity to assemble and/or set up the assistive device. In many cases, this operation is performed only once, at the time of delivery of the assistive device and not by the user himself. This might explain why the content experts did not consider this item as being important. Batavia

and Hammer (1990) also found that "Ease of assembly" (similar to "Installation") finished last on the list of users' priorities for evaluating assistive devices. As for the item "Reaction of others", it refers to the positive and encouraging attitude of others; the content experts did not feel that this item was quite or very important and therefore it could be deleted from the list.

For an instrument to be considered valid its domain must be unambiguously defined and its content must have a generally accepted meaning. At the test level, the content domain of user satisfaction with assistive devices was not challenged in terms of areas to be sampled. At the item level however, very different results were revealed. The experts collectively stressed the need to express more concisely a significant number of the QUEST items and their definitions. It is important to note that this finding was consistent among experts regardless of whether the English, French or Dutch version of QUEST was used. Clearly, the problem of weakly-constructed items was not related to the translation procedures but rather was more fundamentally linked to the item wording and phrasing of the experimental QUEST. Important changes will need to be made to the wording of certain satisfaction items such as "Motivation" or "Device compatibility" in order to ensure optimal content validity.

Another type of evidence of content validity comes from demonstrating that the content is relevant to the purpose of measurement. Experimental QUEST quantifies information about individual users' satisfaction with assistive devices. Its purpose could be to compare clients' satisfaction (description), determine probability of device use (prediction), or assess the outcome of assistive device delivery (evaluation) (Kirshner & Guyatt, 1985). Each alternative has different implications regarding item selection and administration procedures. In the

present study, a majority of experts were interested in assessing satisfaction with specific assistive devices and users' perception and attitudes; they felt it was important to describe individuals within and across groups and to characterize differences of satisfaction.

Case-by-case versus group approaches have clearly influenced the evaluation of administration procedures. On one hand, the interactivity feature of QUEST was viewed positively by the majority of content experts who used QUEST with a small sample of subjects. The tasks themselves (card sorting and sliding bar) were however critiqued in many ways. Simplification and adaptation of the tasks (with non-verbal clients for example) will have to be considered in the final revisions to the assessment instrument. Interestingly, those experts who had tested QUEST on a large number of users, that is the Dutch experts, expressed the need for a short version of QUEST, one that could be administered by mail or by telephone. These suggestions for the development of different QUEST formats for different purposes will be seriously considered in the final revisions of the assessment tool.

The results also revealed that for 50% of the content experts, there was some confusion between the task of rating "importance" and that of rating "satisfaction" of the items. This is an important issue since reliable measures are subjected to a clear understanding of the evaluated variables. The results indicate therefore that if the content is adequately sampled, it may be redundant to ask the users to rate the importance of the QUEST items; the items have already been carefully selected and were considered important. In order to personalize the evaluation and consider individual preferences, it might be more useful to ask the users to identify the three items they consider being the most important. Additional

empirical evidence will however be needed to support this modification before it can be implemented in the new version of QUEST.

Validity is not unchanging nor absolute; quality and quantitative data build on each other to develop an expanding mosaic of validity evidences (Messick, 1980). Following this study with content experts, an alternative method could be employed to gain additional confidence for the results; factorial analysis is the quantitative approach most frequently used in instrument construction to analyze items and their contribution to the test (Nunally, 1978). With the experimental edition of the QUEST, such an analysis would be helpful to confirm correlations between items, eliminate weak relationships or redundancies of a few items and, above all, gain insight into the dimensional structure of the satisfaction concept.

Conclusion

Because cost issues and quality of life is so important in long-term disabling diseases, the validity of outcome measures of assistive technology needs to be thoroughly examined. The results of this study demonstrated that the experimental QUEST was adequately sampled in terms of its content but that some improvements in the wording of the items and in the administration procedures will need to be made. This content validation of QUEST through the use of an international group of content experts has illustrated the many advantages of such partnerships in the development of cross-cultural outcome measures of assistive technology. It is sincerely hoped that this work will serve as an incentive and as a model for uniting educators, researchers and practitioners from different countries in their efforts to develop meaningful and reliable outcome measures in occupational therapy. It is recommended that

future studies establish the reliability and construct validity of QUEST so as to provide a satisfaction assessment tool that will yield scores with appropriate levels of measurement properties for the intended uses in occupational therapy.

Acknowledgments

This paper is based on the doctoral thesis of Louise Demers. The authors thank the content experts from the United States, The Netherlands and Canada who participated in this study. They also gratefully acknowledge the financial support provided by the *Fonds pour la recherche en santé du Québec*, the Canadian Occupational Therapy Foundation and the Foundation of quality and usability research of technical aids in The Netherlands. Special thanks are extended to Mme Francine Giroux, Professor Elisabeth Dutil and Dr Raynald Pineault for their ongoing support.

References

- Andrich, R., Ferrario, M., & Moi, M. (1997). Cost-outcomes analysis in the provision of assistive technology to individual cases. In G. Anagnostis, C. Bühler, & M. Soede (Eds.), *Advancement of assistive technology* (pp. 386-390). Amsterdam, The Netherlands: IOS Press.
- Bain, B. K. (1997). Assistive technology in occupational therapy. In M. E. Neistadt & E. B. Crepeau (Eds.), *Williard and Spackman's occupational therapy* (9th ed., pp. 498-513). Philadelphia, PA: Lippincott.
- Batavia, A. J., & Hammer, G. S. (1990). Toward the development of consumer-based criteria for the evaluation of assistive devices. *Journal of Rehabilitation Research and Development*, 7, 425-436.
- Canadian Association of Occupational Therapy. (1997). *Enabling occupation - An occupational therapy perspective*. Ottawa, ON: CAOT Publications ACE.
- Canadian Burden of Illness Study Group. (1998). Burden of illness of multiple sclerosis: Part 2: Quality of life. *Canadian Journal of Neurologic Sciences*, 25, 31-38.
- Contandriopoulos, A. P., Champagne, F., Potvin, L., Denis, J. L., & Boyle, P. (1990). *Savoir préparer une recherche, la définir, la structurer, la financer*. Montréal, QC: Les Presses de l'Université de Montréal.
- Demers, L., Weiss-Lambrou, R., & Ska, B. (1996). Development of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Assistive Technology*, 8, 3-13.
- DeRuyter, F. (1995) Evaluating outcomes in assistive technology: Do we understand the commitment? *Assistive Technology*, 7, 3-16.
- Finlayson, M., & Havixbeck, K. (1992). A post-discharge study on the use of assistive devices. *Canadian Journal of Occupational Therapy*, 59, 201-207.
- Fisher, A. G., Liu, Y., Velozo, C. A., & Pan, A. W. (1992). Cross-cultural assessment of process skills. *American Journal of Occupational Therapy*, 46, 876-885.

Guillemin, B. B., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology*, *46*, 1417-1432.

Guion, R. M. (1977). Content validity - The source of my discontent. *Applied Psychological Measurement*, *1*(1): 1-10.

Harvey, R. F., Silverstein, B., Venzon, M. A., Kilgore, K. M., Fisher, W. P., Steiner, M. & Harley, J. P. (1992). Applying psychometric criteria to functional assessment in medical rehabilitation: III. Construct validity and predicting level of care. *Archives of Physical Medicine and Rehabilitation*, *73*, 887-892.

Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health and disease. *Journal of Chronic Disease*, *38*, 27-36.

Law, M. (1987). Measurement in occupational therapy: Scientific criteria for evaluation. *Canadian Journal of Occupational Therapy*, *54*, 133-138.

Lorentsen, O. (1997). Goal attainment - A tool to assess outcomes of assistive devices. In G. Anogianakis, C. Bühler, M. Soede (Eds.) *Advancement of Assistive Technology* (pp. 102-104). Amsterdam, The Netherlands: IOS Press.

Mann, W. C., Hurren, D., & Tomita, M. (1993). Comparison of assistive device used and needs of home-based older persons with different impairments. *American Journal of Occupational Therapy*, *47*, 980-987.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, *35*, 1012-1027.

Nunally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Phillips, B., & Zhao, H. (1993). Predictors of assistive technology abandonment. *Assistive Technology*, *5*, 36-45.

Post, K. M. (1993). The promise of assistive technology. *American Journal of Occupational Therapy*, *47*, 965-967.

Reid, D. T., & Jutai, J. (1997). A pilot study of perceived clinical usefulness of a new computer-based tool for assessment of visual perception in occupational therapy practice. *Occupational Therapy International*, *4*, 81-98.

Scherer, M. J., & Vitaliti, L. T. (1997). A functional approach to technological factors and their assessment in rehabilitation. In S. S. Dittmar & G. E. Gresham (Eds.), *Functional assessment and outcome measures for the rehabilitation health professional* (pp. 69-88). Gaithersburg, MD: Aspen.

Smith, R.O. (1996). Measuring the outcomes of assistive technology: Challenge and innovation. *Assistive Technology*, 8, 71-81.

Thorn, D. W., & Deitz, J. C. (1989). Examining content validity through the use of content experts. *Occupational Therapy Journal of Research*, 9, 334-346.

Trachtman, L. (1994). Outcome measures. Are we ready to answer the tough questions? (Editorial). *Assistive Technology*, 6, 91-92.

Valleyrand, R. J. (1989). Vers une méthodologie de validation transculturelle des questionnaires psychologiques: Implications por la recherche en langue française. *Psychologie Canadienne*, 30, 662-680.

Verbrugge, L. M., & Jette, A. M. (1994). The disablement process. *Social Science Medicine*, 38, 1-14.

Vuori, H. (1991). Patient satisfaction - Does it matter? *Quality Assurance in Health Care* ,3, 183-189.

Ward, C. (1990). *Design for all - Consumer needs assessment project year 2 - Results of the second year of a five year study* [Research Report]. Washington, DC: Rehabilitation Engineering Center.

Wessels, R. D., De Witte, L. P., Weiss-Lambrou, R., Demers, L., & Wijnhuizen, G. (1998). A Dutch version of QUEST (D-QUEST) applied as a routine follow-up within the service delivery process. In E. Placencia & E. Ballabio (Eds.), *Improving the Quality of Life for the European Citizen* (pp. 420-424). Washington, DC: IOS Press.

Zastowny, T. R., Roghmann, K. L., & Cafferata, G. L. (1989). Patient satisfaction and the use of health services - Explorations in causality. *Medical Care*, 27, 705-723.

Chapitre 3

Stability and reproducibility of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST)

**Stability and reproducibility of the Quebec User
Evaluation of Satisfaction with assistive
Technology (QUEST)**

Louise Demers, M.Sc., OT(C) *

École de réadaptation, Université de Montréal, Canada
Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Canada

Bernadette Ska, Ph.D

Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Canada

Francine Giroux, M.Sc.

Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Canada

Rhoda Weiss-Lambrou, M.Sc., OT(C)

École de réadaptation, Université de Montréal, Canada

Article submitted to:

Journal of Rehabilitation Outcomes Measurement

* Corresponding author:

Louise Demers

Centre de recherche de l'Institut universitaire de gériatrie de Montréal

4565, Chemin Queen Mary

Montreal (Quebec), Canada, H3W 1W5



Abstract

The Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) is an interview-administered evaluation tool designed to measure satisfaction with assistive technology in a structured and standardized way. The purpose of this study was to determine, at the item level, the test-retest stability and the interrater reproducibility of this instrument in a population of disabled persons using mobility assistive devices. Twenty-four QUEST items were rated on a five-point importance scale and twenty of these same items were scored on a five-point satisfaction scale. A total of 139 subjects were administered the assessment on two occasions, 7 to 11 days apart. For a first cohort of subjects (n=85), the same rater administered both evaluations and for a second cohort (n=54), a different rater was implicated. Reliability coefficients of the satisfaction ratings were found to be moderate to substantial for all the items with regard to stability (weighted Kappa = 0.42 to 0.74) and for 75% of the items in the matter of reproducibility (weighted Kappa = 0.30 to 0.66). The estimated coefficients were perceptibly weaker for the importance ratings (mean Kappa = 0.36) whereas very little variance in scores was observed. Based on the results of this study, the satisfaction scale was reliable in terms of stability and reproducibility however the importance scale did not prove to be reliable for discriminating between and among users of assistive technology.

Keywords: Test-retest stability; Interrater reproducibility; Reliability; Assistive technology; User satisfaction; Outcomes measures.

Background

Outcomes measures have become a critical and challenging issue in rehabilitation (Keith, 1995; Osborn, 1998; Whiteneck, 1994). In the field of assistive technology, outcomes dimensions most typically documented are clinical results, functional status, quality of life, satisfaction and cost (DeRuyter, 1995). All of these indicators play a pivotal role in providing qualitative and quantitative data on the comprehensiveness and effectiveness of assistive technology. Among them, the distinct characteristic of satisfaction measures stems from its unique perspective, which is that of the user. It is a measure that refers to how the client perceives the quality of the services or the product (DeRuyter, 1995). Providers may be good evaluators of objective aspects of assistive technology, however, the real life experience of the users empowers the latter to give the most meaningful judgment of the technology's value and impact. Several researchers (Aharony & Strasser, 1993; Kramer, 1997; Vuori, 1991) consider that a comprehensive and thorough evaluation of a health care experience needs to include clients' appreciation in terms of degree of satisfaction. As a convincing argument, they claim that discrepancies of viewpoints between providers, proxies, and users have been consistently reported in quality evaluation studies (Aharony & Strasser, 1993; Kramer, 1997; Vuori, 1991). Therefore, it is of utmost importance to address each of these groups' points of view separately.

Accountability has been evidenced as a major motivation for assessing satisfaction with assistive technology (DeRuyter, 1995; Trachtman, 1994; Warren, 1993). Indeed, there is an urgent need to demonstrate, especially to the third party payers, its critical value and positive impact on physically disabled

persons' lives. For the assistive technology community, several additional advantages can be expected from this outcome measure. According to Simon and Patrick (1997), measures of satisfaction ideally enable monitoring of the effectiveness of devices provided and services rendered, improving of the services, optimal device utilization, attracting and retaining customers in the context of private sector services, and enhancing positive attitudes among device users. For these reasons, there is general agreement on the need and importance of assessing user satisfaction with assistive technology however, valid and reliable measures are not readily available for use (DeRuyter, 1995; Trachtman, 1994).

To address the need for a satisfaction assessment tool, the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) (Demers, Weiss-Lambrou & Ska, 1996) was developed. It was designed as a structured and standardized measure of satisfaction with a wide range of assistive devices. The concept of satisfaction that guided the construction of QUEST is defined as a person's critical evaluation of several aspects of a device. This evaluation is influenced by one's expectations, perceptions, attitudes and personal values. This definition is consistent with current satisfaction theoretical bases (Keith, 1998; Simon & Patrick, 1997) and was originally inspired by Linder-Pelz' Expectancy-Value model (1982). In order to encompass all the relevant dimensions that influence or are likely to influence satisfaction, the Matching Person and Technology model (Scherer, 1996) was used to structure the instrument's content. Accordingly, three categories of influence were examined: the characteristics of the milieu, the person and the technology itself. Accordingly, the original QUEST is comprised of 24 items related to the technical

characteristics of the device as well as to several psychosocial and service delivery attributes.

In concrete terms, QUEST is divided into three parts and is administered in a face-to-face interview. Part one consists of close-ended questions aimed at describing the context in which user assistive device satisfaction or dissatisfaction developed. In Part two, the evaluator asks the user to rate the degree of importance he or she attributes to the 24 items considered applicable to the assistive device being evaluated. As QUEST was designed to assess a wide range of assistive technology devices, not all items are applicable to every user and every situation. Importance is rated on a 5-point scale, with a score of 1 denoting "of no importance" and 5 indicating "very important". The user is also invited to add any other satisfaction items considered important. In Part three of QUEST, the person is asked to rate his or her satisfaction with 20 of the 24 variables. After pretesting of the tool, four of the 24 items were not assessed for satisfaction because they were found awkward to rate. Each item is scored with a 5-point satisfaction scale, with a score of 1 denoting "not satisfied at all" and 5 indicating "very satisfied". For each item that is scored 3 ("more or less satisfied"), 2 ("not very satisfied") or 1 ("not satisfied at all"), the user is asked to comment or explain the source(s) of dissatisfaction. Finally, the client is asked to score his or her overall satisfaction with the assistive technology device. The instrument was developed simultaneously in French and English and a Dutch version translation was later created (Wessels, De Witte, Weiss-Lambrou, Demers, & Wijlhuizen, 1998).

In studies on patient satisfaction as in other rehabilitation outcomes research, a valid and reliable measure is of critical importance. A valid assessment

instrument is one that measures what is intended to be measured whereas a reliable assessment instrument is one that produces the same result each time it is used, given that the phenomenon addressed is unchanged (Sheikh, 1986). Content validity of QUEST was first examined by a panel of experts during the construction phase of the instrument development (Demers et al., 1996). Additional content validity was recently obtained through the opinions and expertise of 12 assistive technology professionals from different countries (Demers, Wessels, Weiss-Lambrou, Ska, & De Witte, in press). At the test level, this study's findings demonstrated that QUEST is adequately sampled in terms of embracing all the important facets of satisfaction with assistive technology. At the item level however, it was found that changes needed to be made to the wording of the satisfaction variables and to the procedures used in administering QUEST in order to ensure optimal content validity.

The reliability of QUEST for assessing device users' satisfaction has not yet been affirmed. There are three major approaches to the estimation of reliability, depending upon the sources of errors that are considered (Contandriopoulos, Champagne, Potvin, Denis & Boyle, 1990). The first, referred to as test-retest, is concerned with the stability of the measure at two different points in time. According to this approach, subjects are evaluated twice, using the same assessment instrument, and the two sets of scores thereby obtained are compared. The second approach known as interrater reliability assesses the reproducibility of the measure when the phenomenon under study is measured by two or more raters at a given point in time. In the case of interview format measures such as QUEST, interrater and intertime reliability are imbricated - what is estimated is a person's score on QUEST when administered by two

different raters at two points in time (Feinstein, 1987). Similarly, two sets of scores are collected and correlated to produce a reliability coefficient. The third approach is referred to as the internal consistency. Its focus is on the extent to which items of a composite score substantively measure the same concept. It is a measurement property that characterizes the scale or the subscales in their entirety. Stability and reproducibility estimators can be computed separately for each item as well as for the whole scale if its unidimensional structure has been adequately demonstrated (Silverstein, Kilgore, Fisher, Harley & Harvey, 1991).

The QUEST instrument was not originally developed to yield a total score since the dimensional structure of the satisfaction construct has yet to be established. In previous applied studies using QUEST, results at the item level were examined to track patient satisfaction and to detect specific areas that are problematic and dissatisfying (Quintin, Weiss-Lambrou & McCabe, 1997; Weiss-Lambrou, Demers, Tremblay, Ska, Lacoste & Dansereau, 1998; Wessels et al., 1998). Other ongoing studies aim at describing users' degree of satisfaction with specific characteristics of assistive devices and therefore, item level measurement properties need to be established. The purpose of the present study was to measure at the item level, the test-retest stability and the interrater reproducibility of QUEST with adult users of different types of mobility devices. This type of assistive device was selected because of its widespread use by persons with disabilities.

Methods

Subjects

To be included in the study, subjects were required to meet the following two selection criteria: (1) they had used an assistive device for at least four months; and (2) the technology used was a mobility device, either a wheelchair (manual or powered driven, with or without postural aids) or a lower limb prosthesis. Subjects who met the above criteria were excluded from the study if: (a) they presented major sensory deficiency of vision and/or hearing; (b) they had a cognitive impairment as evidenced by a score less than 25/30 (83.3%) on the Mini-Mental State Examination (MMSE) (Braekhus, Laake & Engedal, 1992; Folstein, Folstein & McHugh, 1975), and/or (c) they did not speak French fluently.

One hundred and fifty-eight adult device users subjects (82 men and 76 women) were recruited from the assistive technology programs located in three Montreal rehabilitation facilities: *Centre de réadaptation Lucie-Bruneau*, *Institut de réadaptation de Montréal*, and *Centre de réadaptation Constance Lethbridge*. In two of these facilities, the eligible candidates were identified through an ongoing repairs and servicing appointments registry. In the third setting, the assistive technology program charts were reviewed by a research assistant. A first cohort of subjects was assigned to the test-retest study (n=97) and a second cohort to the interrater study (n=61). All participants provided informed consent. Because of some health problems and other unexpected events and circumstances, 12 subjects were not available for the second testing situation. In the time interval between the testing situations, an additional six subjects had experienced certain changes, either personally or with regard to their device. Because their

satisfaction ratings could be affected by these changes, they were removed from the sample. Consequently, 85 subjects remained for the test-retest study and 54 for the interrater study.

Based on the data gathered from Part one of QUEST, the characteristics of the samples in the test-retest study and the interrater study are displayed in Table I. The first layer of the table presents the general characteristics of the device users. In both studies, the mean age for the subjects was 51 years old. The percentage of male subjects was 52.9% in the test-retest study, compared with 44.4% for the interrater study. More than half the subjects reported to have participated in the selection of their device. The percentage appeared to be somewhat higher in the test-retest study (72.6%) than in the interrater study (51.9%). Also, the number of functional problems reported is slightly different between the two groups, respectively 2.7 and 4.0 functional problems.

General information characterizing the devices studied are grouped in the second layer of Table I. There was a slightly greater proportion of wheelchair users in the interrater group, 88.9%, compared with 80.0 % for the test-retest group. More than 80% of the subjects used their devices at least everyday and approximately 60% of the sample reported previous experience with one or more devices of the same kind. The mean length of time since the device was obtained was 42 months (3.5 years).

Table I: Subjects, devices and environment characteristics of the two samples

		Test-retest study (n=85)	Interrater study (n=54)
Users	Mean age (years) Range	50.7 ± 13.4 21.5 - 79.4	50.6 ± 13.95 22 - 83
	Gender (% male)	45/85 (52.9%)	24/54 (44.4%)
	Mean MMSE score (%) Range	97.5 ± 3.3 83.3 - 100	97.1 ± 3.5 83.3 - 100
	Participation in selection of device (% yes)	61/84 (72.6%)	28/54 (51.9%)
	Mean number of functional problem areas Range	2.7 ± 2.2 0 - 8	4.0 ± 2.3 0 - 8
Assistive device	Number of wheelchairs (% total)	68/85 (80.0%)	48/54 (88.9%)
	Number of prothesis (% total)	20/85 (20.0%)	6/54 (11.1%)
	Frequency of device use (% everyday and always)	69/85 (81.2%)	47/54 (87.0%)
	Previous experience with other devices of the same kind (% yes)	33/57 (57.9%)	24/39 (61.5%)
	Mean length of time since device was obtained (months) Range	40.2 ± 28.2 4 - 120	43.3 ± 28.2 5 - 121
Environment	Living situation (% living alone)	41/84 (48.8%)	32/53 (60.4%)
	Home-based services (% yes)	55/84 (65.5%)	34/53 (64.2%)
	Provider of assistive device delivery services (% rehabilitation unit)	77/83 (92.8%)	50/52 (92.6%)
	Training (% yes)	24/83 (28.9%)	16/54 (29.6%)
	Waiting period for device delivery (weeks) Range	18.0 ± 13.1 0 - 52	25.8 ± 21.1 0 - 103
	Funding source (% government insurance)	83/85 (97.6%)	50/54 (92.6%)

Legend: MMSE Mini-Mental State Examination

The last grouping of general information describes the environmental characteristics of the subjects for both samples. A substantial proportion (60.4%) of the subjects assigned to the interrater study lived alone while the majority of subjects assessed in the test-retest study shared their housing with a spouse, relatives or other companions (51.2%). A similar finding was found for the percentage of subjects who received home-based services (65%). Almost all subjects had obtained their device from an assistive technology program located in a rehabilitation facility. Assistive device training was provided to approximately 29% of the subjects. With regard to the mean waiting period for device delivery, some slight discrepancies between the two groups were present (18.0 and 25.8 months). More than 92% of the subjects had their devices financed by the Quebec government health insurance program.

Raters

In the test-retest study, subjects were assigned to one of three participating raters. Because of differences in rater availability, the number of subjects interviewed by each rater was 42, 22 and 16 subjects. The remaining 5 subjects were tested by the student researcher. In the interrater study, three raters (two of whom were the same as for the test-retest study) were assigned in random pairs to individual subjects. Respectively 19, 18, and 16 subjects were interviewed by each twosome. In terms of professional background, the three raters had a postgraduate training in biomedical research. They received a six-hour training session that focused mainly on the procedures used to administer QUEST. Written instructions were provided and simulated interviews and pre-testing were conducted to ensure maximum consistency between raters.

Procedure

The same procedure was applied to the test-retest and the interrater studies. Selected subjects were first contacted by telephone to confirm their willingness to participate in the study and to schedule the two evaluation appointments. The interviews were conducted in the participants' home and included, for the first meeting, an explanation of the purpose of the study, signing the informed consent form, completing the MMSE and the administration of QUEST. This first meeting lasted 35 to 75 minutes, depending on the number of sources of dissatisfaction that were reported and explored. The second meeting consisted of administering QUEST Part two and Part three. Consequently, less time was required, approximately 15 to 30 minutes. A critical problem in assessing degree of test-retest stability and interrater reproducibility at two time points is determining a time interval long enough to minimize the potential confounding influence of memory but yet short enough to reduce the possibility of real changes in satisfaction (Crocker & Algina, 1986; Pedhazur & Schmelkin, 1991). An interval of 7 to 11 days was considered reasonable for QUEST and was chosen for this investigation.

Statistical Analysis

Analysis were carried out on an IBM PC, using SPSS 8.0 and Excel 97 softwares. Descriptive statistics such as means and standard deviation were computed. Stability and reproducibility were assessed by calculating Kappa (Cohen, 1960) and weighted Kappa (Cohen, 1968) for each of the items, that is, 20 items out of 24 items for the satisfaction scale and the complete item pool for the importance scale. The Kappa and weighted Kappa statistics are preferred to the simpler measure of percentage of agreement because they discount the

proportion of agreement that is expected by chance alone (Maclure & Willett, 1987). With ordinal data, weighted Kappa is an appropriate choice because it takes into account partial agreement. Accordingly, a pair of observations two categories apart is regarded as a greater disagreement than a pair in adjacent categories. In this study, linear agreement weights recommended by Cicchetti (1976) and the revised standard error formula given by Fleiss, Cohen and Everitt (1969) were applied. To compute Kappa and weighted Kappa statistics, the data structure required is a square $k \times k$ table, thus implying that the scope of response options be the same on the two administrations of QUEST. If this criterion was not met, an alternate estimator was computed by combining QUEST scores of 1, 2 and 3 into a single category, and scores 4 and 5 in a separate category.

Reliability estimates were reported for the pooled data. According to Cicchetti and Fleiss (cited in Cicchetti, 1976), a minimal sample size of 50 subjects is required to compute weighted Kappa using a five-point rating scale. Consequently, for individual raters and pairs of raters, the divided samples would be insufficient to produce valid results. Judgments on the strength of agreement were based on Landis and Koch's (1977) benchmarks for categorical data. Hence, coefficients values were interpreted according to the following categories: poor (< 0.00), slight ($0.00 - 0.20$), fair ($0.21 - 0.40$), moderate ($0.41 - 0.60$), substantial ($0.61 - 0.80$), and almost perfect ($0.81 - 1.00$). Good performing items were expected to demonstrate a moderate to higher level of agreement.

Results

Stability and reproducibility of importance ratings

For the rating of importance, the descriptive statistics such as means and standard deviations were very similar between the two sets of observations, both for the test-retest and the interrater studies. Therefore, only the QUEST data gathered during the first session are reported in Table II. For most of the items, the number of subjects evaluated came very close to or exceeded the 50 required subjects. As evidenced in Table II (under column "n"), exceptions were found mostly in the interrater study for items *installation* (#4), *cost* (#6), and *device compatibility* (#16). Close inspection of the table also reveals that in both data sets a large number of items presented severe positively skewed distributions. The mean scores of importance averaged 4.34 and standard deviations lower than 1.00 were obtained in 66.7% of the cases. Some variability in subjects scores were however found for the items *cost* (#6), *appearance* (#14), *device compatibility* (#16), *weight* (#17), *training* (# 20), *social circle support* (#22), and *reaction of others* (#23). For many items, the subjects' responses did not cover the full range of possible scores; for some, the scores clustered in the upper range of the scale and, for others, the number of chosen response options was inconsistent between the two sets of observations. As a result, it was not possible in either study to compute the weighted Kappa statistics for 15 of the 24 variables (62.5%). For this reason, an alternate reliability estimator (Kappa) was computed by combining importance scores of 1, 2 and 3 into a single category, and scores of 4 and 5 into a separate category.

Table II: Descriptive statistics of QUEST importance and satisfaction ratings

Item	Importance ratings				Satisfaction ratings			
	Test-retest Study (n=85)		Interrater Study (n=-54)		Test-retest Study (n=85)		Interrater Study (n=-54)	
	n	Mean & SD	n	Mean & SD	n	Mean & SD	n	Mean & SD
1. Simplicity of use	85	4.79 ± 0.44	54	4.76 ± 0.43	85	4.15 ± 0.91	53	4.02 ± 1.09
2. Repairs/servicing	85	4.69 ± 0.62	50	4.89 ± 0.38	82	3.98 ± 1.08	48	3.96 ± 1.21
3. Maintenance	83	4.47 ± 0.74	50	4.29 ± 0.91	80	4.24 ± 0.86	43	3.98 ± 0.98
4. Installation	75	4.48 ± 0.82	41	4.47 ± 0.63	74	4.25 ± 0.90	34	3.95 ± 1.05
5. Effectiveness	85	4.80 ± 0.53	53	4.81 ± 0.44	85	4.25 ± 0.87	53	4.22 ± 0.79
6. Cost	45	3.14 ± 1.41	27	3.14 ± 1.61	30	4.46 ± 0.88	15	4.38 ± 1.06
7. Professional service	84	4.62 ± 0.64	52	4.56 ± 0.72	85	4.02 ± 1.00	48	4.00 ± 1.22
8. Durability	84	4.70 ± 0.58	53	4.75 ± 0.52	77	4.19 ± 0.89	50	4.17 ± 0.98
9. Multi-purposefulness	82	4.45 ± 0.76	50	4.28 ± 0.99	78	4.06 ± 1.08	48	3.96 ± 1.18
10. Adjustments	84	4.63 ± 0.64	51	4.62 ± 0.71	81	3.94 ± 0.95	47	3.73 ± 1.07
11. Comfort	85	4.85 ± 0.45	54	4.89 ± 0.37	85	3.90 ± 0.97	51	3.92 ± 1.17
12. Service delivery	83	4.43 ± 0.65	51	4.35 ± 0.87	83	3.91 ± 0.97	51	3.52 ± 1.38
13. Follow-up services	83	4.42 ± 0.73	50	4.43 ± 0.81	80	4.04 ± 1.04	45	3.76 ± 1.29
14. Appearance	85	3.89 ± 1.20	53	3.67 ± 1.18	83	4.18 ± 0.84	51	4.17 ± 0.85
15. Transportation	80	4.66 ± 0.59	50	4.63 ± 0.66	78	4.11 ± 0.95	50	3.94 ± 1.19
16. Device compatibility	54	4.29 ± 0.91	30	4.03 ± 1.16	51	4.16 ± 0.90	23	4.29 ± 0.68
17. Weight	84	3.93 ± 1.40	50	4.13 ± 1.11	74	3.87 ± 1.08	45	3.68 ± 1.22
18. Safety	83	4.78 ± 0.54	53	4.67 ± 0.67	83	4.24 ± 0.89	53	4.02 ± 0.94
19. Dimensions	85	4.52 ± 0.75	53	4.57 ± 0.81	85	3.94 ± 1.13	53	3.96 ± 1.18
20. Training	77	3.84 ± 1.15	46	3.80 ± 1.31	55	3.89 ± 1.28	29	3.78 ± 1.36
21. Motivation	77	4.54 ± 0.91	50	4.79 ± 0.54				
22. Social circle support	78	3.83 ± 1.38	48	4.10 ± 1.20				
23. Reaction of others	80	3.04 ± 1.37	51	2.83 ± 1.55				
24. Effort	79	4.28 ± 1.09	49	4.63 ± 0.72				

Legend: SD Standard Deviation.

Table III presents the stability and reproducibility coefficients of the 24 QUEST items importance ratings. General appraisal of the results displayed for stability reveals that the estimated stability coefficients are perceptibly weaker than expected. A substantial number of items (n=13, 54.2%) demonstrated fair or poor degrees of agreement and consequently did not reach an acceptable level of reliability. These items were *simplicity of use* (#1), *repairs/servicing* (#2), *maintenance* (#3), *installation* (#4), *effectiveness* (#5), *professional services* (#7), *durability* (#8), *multipurposefulness* (#9), *adjustments* (#10), *transportation* (#15), *device compatibility* (#16), *safety* (#18), and *dimensions* (#19). Low variability in responses could partly explain these results since the items that demonstrated moderate (n=8, 33.3%) and substantial agreement (n=3, 12.5%) also tended to have less skewed distributions (7 of 11 items). Items found in this category were *cost* (#6), *comfort* (#11), *service delivery* (#12), *follow-up services* (#13), *appearance* (#14), *weight* (#17), *motivation*, (#20), *training* (#21), *social circle support* (#22), *reaction of others* (#23), and *effort* (#24).

In terms of reproducibility, the coefficients for the items *simplicity of use* (#1) and *comfort* (#11) could not be estimated neither with weighted Kappa nor with Kappa, because of severely reduced scope of responses. Of the 22 testable items, 14 items (63.6%) were judged as poor (n=8, 36.4%) or fair (n=6, 27.3%) and did not make the cutoff for acceptable reliability. As expected, most items (n=18, 78.3%) either maintained the same level as in the test-retest study (n=7) or dropped down either to the adjacent category (n=7) or to two categories below (n=1). However, seven items (29.2%) demonstrated a rather different pattern of responses by showing increased reliability. These results might be attributed to inappropriate frequency distributions rather than to real improvement.

Table III: Coefficients of stability and reproducibility of 24 QUEST items:
Importance ratings

Item	Stability		Reproducibility	
	K	Judgment	K	Judgment
1. Simplicity of use	-0.01	Poor	UN	
2. Repairs/servicing	-0.02	Poor	-0.02	Poor
3. Maintenance	0.09	Poor	0.39**	Fair
4. Installation	0.31**	Fair	0.29*	Fair
5. Effectiveness	-0.02	Poor	-0.02	Poor
6. Cost	0.73**	Substantial	0.54**	Moderate
7. Professional service	0.36**	Fair	-0.06	Poor
8. Durability	-0.02	Poor	1.00**	Almost perfect
9. Multi-purposefulness	0.13	Poor	0.17	Poor
10. Adjustments	0.31**	Fair	-0.08	Poor
11. Comfort	0.49**	Moderate	UN	
12. Service delivery	0.45**	Moderate	0.49**	Moderate
13. Follow-up services	0.47**	Moderate	0.13	Poor
14. Appearance	0.61**	Substantial	0.84**	Almost perfect
15. Transportation	0.21	Poor	0.23	Poor
16. Device compatibility	0.27*	Fair	0.34*	Fair
17. Weight	0.61**	Substantial	0.49**	Moderate
18. Safety	0.38**	Fair	0.65**	Substantial
19. Dimensions	0.08	Poor	0.30*	Fair
20. Training	0.57**	Moderate	0.36**	Fair
21. Motivation	0.41**	Moderate	0.38**	Fair
22. Social circle support	0.53**	Moderate	0.78**	Substantial
23. Reaction of others	0.50**	Moderate	0.68**	Substantial
24. Effort	0.48**	Moderate	0.15	Poor

Legend: **: $p < 0.01$; *: $p < 0.05$; UN: Untestable; K: Kappa.

Stability and reproducibility of satisfaction ratings

In terms of the ratings of satisfaction the means and standard deviations were very similar between the two sets of observations. Consequently, statistics reported in Table II were based on data gathered during the first meeting with the subjects. Because of the non-applicable response option, the samples sizes fluctuated according to individual items, from 30 to 85 subjects for the test-retest study and from 15 to 53 subjects for the interrater study. It should be noted that in some cases, the number of subjects fell much lower than the recommended 50 subjects in the interrater study. The most seriously affected variables were *installation* (#4), *cost* (#6), *device compatibility* (#16), and *training* (#20). Satisfaction distributions were positively but not severely skewed, with scale means ranging from 3.87 (#17, *weight*) to 4.46 (#6, *cost*) for the test-retest study, and from 3.52 (#12, *service delivery*) to 4.38 (#6, *cost*) for the interrater study. Concurrent variability in responses tended to be modest, with standard deviations scattered between 0.84 (#14, *appearance*) and 1.28 (#21, *training*) for the test-retest study, and between 0.68 (#16, *device compatibility*) and 1.29 (#13, *follow-up services*) for the interrater study. Individual subject ratings however covered the full range of possible scores with a minimum value at 1.00 and a maximum value at 5.00. An exception was found for *cost* (#6) for which the lowest value obtained was 2.00.

The coefficients of stability and reproducibility derived from this data are displayed for the satisfaction ratings in Table IV. Inspection of the statistics listed under "Stability" indicate that the entire set of stability coefficients reached the moderate and substantial agreement levels defined by Landis & Koch (1977). The most stable items were *repairs/servicing* (#2), *effectiveness* (#5),

Table IV: Coefficients of stability and reproducibility of 20 QUEST Items:
Satisfaction ratings

Item	Stability		Reproducibility	
	WK	Judgment	WK	Judgment
1. Simplicity of use	0.52**	Moderate	0.35**	Fair
2. Repairs/servicing	0.68**	Substantial	0.65**	Substantial
3. Maintenance	0.42**	Moderate	0.33**	Fair
4. Installation	0.43**	Moderate	0.49**	Moderate
5. Effectiveness	0.62**	Substantial	0.55**	Moderate
6. Cost	0.55**	Moderate	0.46**	Moderate
7. Professional service	0.61**	Substantial	0.48**	Moderate
8. Durability	0.69**	Substantial	0.40**	Moderate
9. Multi-purposefulness	0.58**	Moderate	0.30*	Fair
10. Adjustments	0.52**	Moderate	0.43**	Fair
11. Comfort	0.51**	Moderate	0.72**	Substantial
12. Service delivery	0.61**	Substantial	0.51**	Moderate
13. Follow-up services	0.56**	Moderate	0.63**	Substantial
14. Appearance	0.48**	Moderate	0.53**	Moderate
15. Transportation	0.42**	Moderate	0.47**	Moderate
16. Device compatibility	0.61**	Substantial	0.36**	Fair
17. Weight	0.63**	Substantial	0.69**	Substantial
18. Safety	0.58**	Moderate	0.43**	Moderate
19. Dimensions	0.74**	Substantial	0.66**	Substantial
20. Training	0.62**	Substantial	0.50**	Moderate

Legend: ** : $p < 0.01$; * : $p < 0.05$; WK: Weighted kappa.

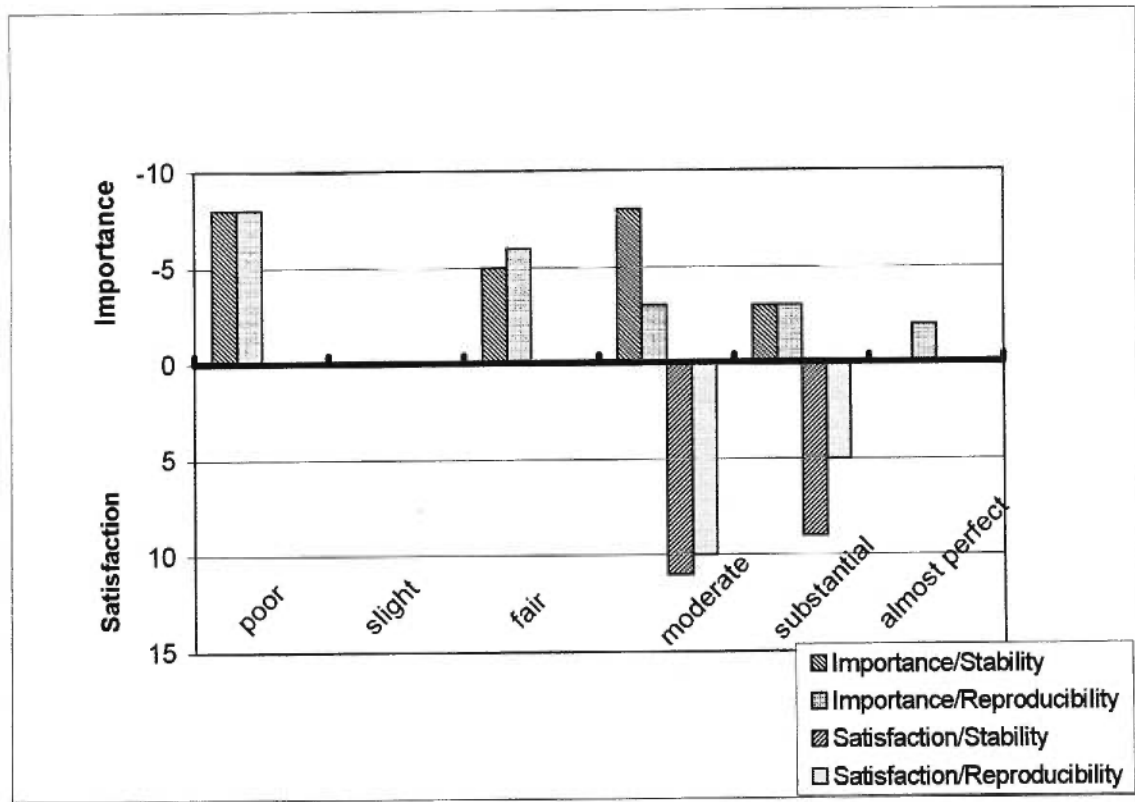
professional services (#7), *durability* (#8), *service delivery* (#12), *device compatibility* (#16), *weight* (#17), *dimensions* (#19), and *training* (#20).

With reference to reproducibility, results reveal that most items either maintained the same agreement level as in the test-retest study (n=11, 55.0%) or dropped to the next category (n=6, 30.0%). Only one item (#16, *device compatibility*) fell two category levels below (5.0%). Such results are not surprising since an additional source of error was imposed on the research design, one that is attributed to dissimilarities between raters. Quite unexpectedly, two items increased their reliability level from moderate to substantial, namely *comfort* (#11) and *follow-up services* (#14). In total, 15 items (75.0%) reached the expected level of reproducibility. These were: *repairs and servicing* (#2), *installation* (#4), *effectiveness* (#5), *cost* (#6), *professional services* (#7), *durability* (#8), *comfort* (#11), *service delivery* (#12), *follow-up services* (#13), *appearance* (#14), *transportation* (#15), *weight* (#17), *safety* (#18), *dimensions* (#19), and *training* (# 21). On the other hand, five items (25.0%) did not meet the established criteria and fell into the category of “fair” items. They included *simplicity of use* (#1), *maintenance* (#3), *multi-purposefulness* (# 9), *adjustments* (#10), and *device compatibility* (#16).

To graphically summarize the results of this paper, Figure 1 illustrates the frequency distribution of the items in terms of the levels of agreement as defined by Landis and Koch (1977). This figure incorporates results for stability and reproducibility, on both the importance scale and the satisfaction scale. As shown by the bars, reliability of the items scored on the importance scale tends to fluctuate and a substantial proportion of items are rated as “poor” and “fair”. When items are scored on the satisfaction scale, reliability is more consistent

throughout the items, at “moderate” and “substantial” levels of agreement. These results conclude that the importance scale proved to be unreliable but that the satisfaction scale is reliable in terms of stability and reproducibility.

Figure 1: Frequency distributions of items per category of agreement (n=6)



Discussion

The concept of reliability reflects the extent to which the amount of error, either random or systematic, is present in an assessment instrument. The appraisal of this measurement property was essential to establishing how well QUEST can differentiate among and between individuals. Indeed, the error of measurement ideally accounts for a very small fraction of the variability observed in a data set. In this study, both the importance and satisfaction ratings were submitted to two procedures for estimating reliability. The first pertained to the stability of the measure, and the second focused on interrater reproducibility at two time points. Based on previous and intended uses and applications of QUEST, these procedures were felt to be most appropriate.

Results revealed that the QUEST items demonstrate acceptable levels of reliability for the satisfaction ratings. All computed coefficients were significantly different from zero, indicating that some agreement beyond chance occurred in all cases. Based on Landis and Koch (1977), the strength of this agreement following test-retest was judged either substantial or moderate, for the entire 20 items pool. Interrater reproducibility at two time points also met the threshold of acceptability for 75% of the item content. The remaining five items (*simplicity of use, maintenance, multipurposefulness, adjustments, and device compatibility*) were however considered only as fairly reliable. In order to increase the estimated coefficients, one option would be to examine and improve item formulations. More straightforward wording could reduce interpretative error induction in the administration of QUEST.

One important recommendation reported in a recent study on the QUEST content validity pertained to the clarity of the items (Demers et al., in press). Because some items were ambiguous or difficult to explain, changes needed to be made to simplify the phrasing or wording, such as in the case of *multipurposefulness* and *device compatibility*. The present findings are somewhat at variance with this conclusion since most of the original items nevertheless performed well in terms of reliability of satisfaction ratings. By reducing error variance, future wording clarifications are likely to improve the items reliability estimates. Furthermore, considering the relatively unstable nature of satisfaction, the results of the present study can and should be viewed positively.

Substantial differences in the degree of reliability were found for the importance ratings. The purpose of Part two of QUEST was to personalize the evaluation in terms of individual device users' preferences. This objective was not met because almost all the items were considered as being very important. Indeed for a substantial number of items ($n=17$), subjects' scores averaged close to or higher than 4.50/5.00. Concurrently, the corresponding standard deviations reflecting data dispersion resulted in very low values. From a validity perspective, these results confirmed that QUEST had adequate content coverage, by embracing those aspects of satisfaction that were considered important to the users. However in terms of reliability, the obtained coefficients did not support the usefulness of rating the importance of each of the QUEST items. More than half the items did not generate a moderate to higher level of agreement. Furthermore, eight items rated poorly for test-retest stability as well as interrater reproducibility. It was shown that the amplitude of the reliability

coefficients is directly related to the variability between subjects (Crocker & Algina, 1986). Therefore, the lack of data variance in the importance scores can explain most of these results. It was therefore decided that the coefficients were not strong enough to provide an accurate representation of users' preferences. As a result, it was found necessary to withdraw Part two from QUEST; the scoring for the importance items did not prove to be reliable for discriminating between and among users of assistive technology.

A limitation of this study that needs to be raised pertains to the sample size of both the stability and the reproducibility studies. For the pooled data, the number of participating subjects was sufficient to produce significant reliability coefficients ($p < 0.01$). For individual raters and pairs of raters however, the splitted samples would not have reached the minimal required size of 50 subjects and unstable estimators of reliability would have ruled out valid comparisons between raters and pairs of raters. Consequently, errors that could have been attributed to an individual rater or to a pair of raters may have gone undetected and possibly contaminated the data sets. Also, the generalizability of the results are restricted to the population from which these studies' subjects were recruited. From this perspective, the 20 QUEST items proved to be reliable for assessing item satisfaction with users of wheelchairs and lower limb prostheses. The results of this study need to be interpreted with caution before applying them with other types of assistive devices users.

In conclusion, this study's findings firmly support the use of QUEST as a reliable instrument for assessing users' satisfaction with assistive technology. Most of the item content proved to perform very well, with regards to both the test-retest stability and the interrater reproducibility. These results are considered very

positive in that QUEST can now serve as a valid and reliable outcomes measures in the field of assistive technology. Administration of the instrument assessment will be particularly relevant in studies whereby changes of assistive device characteristics are of interest. The results from the importance ratings also revealed very meaningful information. The fact that a substantial proportion of the items were considered of great importance to the subjects, with little discordant viewpoints, is noteworthy. Indeed, discriminating among subjects did not seem to be that critical because the most important items had already been selected during the construction phase of the instrument. As a result, the revised version of QUEST will not include the rating of item importance. Future QUEST studies are recommended to improve specific items formulations, to explore the dimensional structure of the satisfaction concept, and to evaluate the reliability of the scale and subscales as a whole.

Acknowledgments

This paper is part of the doctoral thesis of Louise Demers. The authors gratefully acknowledge financial support from the *Fonds pour la recherche en santé du Québec*, the Canadian Occupational Therapy Foundation and *the Université de Montréal*. They are also indebted to the assistive technology users and the personnel from the *Centre de réadaptation Lucie-Bruneau*, *Institut de réadaptation de Montréal* et *Centre de réadaptation Constance Lethbridge*. Special thanks to Camille Tremblay, Michèle Monette, Brigitte Vachon, and Annick Bouliane for their invaluable assistance in the data collection.

References

- Aharony, L., & Strasser, D. (1993). Patient satisfaction: What we know and what we still need to explore. *Medical Care Review, 50*, 49-79.
- Braekhus, A., Laake, K., & Engedal, K. (1992). The mini-mental state examination: Identifying the most efficient variables for detecting cognitive impairment in the elderly. *Journal of American Geriatrics Society, 40*, 1139-1145.
- Cicchetti, D. V. (1976). Assessing inter-rater reliability for rating scales: Resolving some basic issues. *British Journal of Psychiatry, 129*, 452-456.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*, 213-220.
- Contandriopoulos, A. P., Champagne, F., Potvin, L., Denis, J. L., & Boyle, P. (1990). *Savoir préparer une recherche, la définir, la structurer, la financer*. Montréal, QC: Les Presses de l'Université de Montréal.
- Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Orlando, FL: Harcourt Brace Jovanovich.
- Demers, L., Weiss-Lambrou, R., & Ska, B. (1996). Development of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Assistive Technology, 8*, 3-13.
- Demers, L., Wessels, R., Weiss-Lambrou, R., Ska, R., & De Witte, L. (in press). An international content validation of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Occupational Therapy International*.
- DeRuyter, F. (1995) Evaluating outcomes in assistive technology: Do we understand the commitment? *Assistive Technology, 7*, 3-16.
- Feinstein, A. R. (1987). *Clinimetrics*. New Haven, CT: Yale University Press.
- Fleiss, J. L., Cohen, J., & Everitt, B. S. (1969). Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin, 72*, 323-327.

- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatry Research*, *12*, 189-198.
- Keith, R. A. (1995). Conceptual basis of outcome measures. *American Journal of Physical Medicine and Rehabilitation*, *74*, 73-80.
- Keith, R.A. (1998). Patient satisfaction and rehabilitation services. *Archives of Physical Medicine and Rehabilitation*, *79*, 1122-1128.
- Kramer, A. M. (1997). Rehabilitation care and outcomes from the patient's perspective. *Medical Care*, *35*, js48-js57.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.
- Linder-Pelz, S. (1982). Toward a theory of patient satisfaction. *Social Science and Medicine*, *16*, 577-582.
- Maclure, M., & Willett, W. (1987). Misinterpretation and misuse of the kappa statistic. *American Journal of Epidemiology*, *126*, 161-169.
- Osborn, C. E. (1998). Developing instruments for assessment of patient outcomes. *Journal of Rehabilitation Outcomes Measurement*, *2*(6), 18-25.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Quintin, S., Weiss-Lambrou, R., & McCabe, D. (1997). *Enquête sur l'utilisation et la satisfaction concernant le lève-personne sur rail au plafond utilisé à domicile* [Survey of utilization and satisfaction with homebased used lifts] (Research Report). Montreal, QC, Canada: Société Logique.
- Scherer, M. J. (1996). *Living in the state of stuck – How technology impacts the lives of people with disabilities* (2d ed.). Cambridge, MA: Brookline.
- Sheikh, K. (1986). Disability scales: Assessment of reliability. *Archives of Physical Medicine and Rehabilitation*, *67*, 245-249.

- Silverstein, B., Kilgore, K. M., Fisher, W. P., Harley, J. P., & Harvey, R. F. (1991). Applying psychometric criteria to functional assessment in medical rehabilitation: I. Exploring unidimensionality. *Archives of Physical Medicine and Rehabilitation*, 72, 631-637.
- Simon, S. E., & Patrick, A. (1997). Understanding and assessing consumer satisfaction in rehabilitation. *Journal of Rehabilitation Outcomes Measurement*, 1(5), 1-14.
- Trachtman, L. (1994). Outcome measures. Are we ready to answer the tough questions? (Editorial). *Assistive Technology*, 6, 91-92.
- Vuori, H. (1991). Patient satisfaction - Does it matter? *Quality Assurance in Health Care*, 3, 183-189.
- Warren, C. G. (1993). Cost effectiveness and efficiency in assistive technology service delivery. *Assistive Technology*, 5, 61-65.
- Weiss-Lambrou, R., Demers, L., Tremblay, C., Ska, B., Lacoste, & Dansereau, J. (1997). In QUEST of user satisfaction with assistive technology devices. In G. Anogianakis, C. Bülher, & M. Soede (Eds.), *Advancement of Assistive Technology* (pp. 428-431). Washington, DC: IOS Press.
- Wessels, R. D., De Witte, L. P., Weiss-Lambrou, R., Demers, L., & Wijnhuizen, G. (1998). A Dutch version of QUEST (D-QUEST) applied as a routine follow-up within the service delivery process. In E. Placencia & E. Ballabio (Eds.), *Improving the Quality of Life for the European Citizen* (pp. 420-424). Washington, DC: IOS Press.
- Whiteneck, G. G. (1994). Measuring what matters: Key rehabilitation outcomes. *Archives of Physical Medicine and Rehabilitation*, 75, 1073-1076.

Chapitre 4

Item analysis of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST)

**Item analysis of the Quebec User Evaluation of
Satisfaction with assistive Technology (QUEST)**

Louise Demers, M.Sc., OT(C) *

École de réadaptation, Université de Montréal, Canada
Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Canada

Rhoda Weiss-Lambrou, M.Sc., OT(C)

École de réadaptation, Université de Montréal, Canada

Bernadette Ska, Ph.D

Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Canada

Article submitted to:
Assistive Technology

*Corresponding author:

Louise Demers

Centre de recherche de l'Institut universitaire de gériatrie de Montréal

4565 Chemin Queen Mary

Montreal (Quebec), Canada, H3W 1W5



Abstract

The Quebec User Evaluation of Satisfaction with assistive Technology (QUEST) is an outcomes assessment tool designed to measure satisfaction with assistive technology in a structured and standardized way. The purpose of this paper is to present the results of the item analysis of the 24 items comprising QUEST and to explain how a subset of items demonstrating optimal measurement performance was selected. The criteria against which the items were measured were: general acceptability, content validity, criterion validity, contribution to internal consistency, test-retest stability, and sensitivity. The items that ranked best in terms of these measurement properties were submitted to factorial analysis in order to complete the item selection. The first series of analysis reduced the item pool approximately by half and the second series of analysis led to the final selection of 12 items. Factor analysis results suggested that the underlying structure of satisfaction with assistive technology consists of two dimensions related to the assistive technology DEVICE (eight items) and SERVICES (four items). This finding was cross-validated in a different sample of subjects and an identical structure was obtained thereby strongly supporting the stability of the satisfaction model. The 12 item revised version that will result from this study should prove to be a reliable and valid instrument for measuring outcomes in the field of assistive technology.

Keywords: Assistive technology; Item analysis; User satisfaction; Instrument development; Reliability; Validity.

Introduction

Measuring satisfaction is conferred considerable importance in the management of assistive technology device outcomes (DeRuyter, 1995; Kohn, Mortola & LeBlanc, 1991; Trachtman, 1994). Its positive influences encompass enabling the monitoring of services, quality improvement, optimal device utilization, attraction and retention of customers, and positive attitudes among users (Simon & Patrick, 1997). To be useful as a gauge of user opinion, a satisfaction measure should meet some basic criteria: it needs to be reliable and valid, appropriate for use with a wide range of assistive devices, and sufficiently sensitive to differentiate among individuals in terms of their degree of satisfaction. In the context of instrument development, such a tool needs to evolve over time and improve itself through field-testings and controlled studies.

All too frequently in the past, investigators have tried to assess satisfaction with a general unidimensional factor presumed to represent overall satisfaction with devices. However, several evidences derived from the health care domain have demonstrated the relevance of a multidimensional concept of satisfaction (Hall & Dornan, 1988; Rubin, 1990; Ware, Davies-Avery & Stewart, 1978). Such a framework is believed to be useful as well in the field of assistive technology. For this reason and because no such instrument was readily available, the Quebec User Satisfaction with assistive Technology (QUEST) was created as a structured and standardized measure of user satisfaction with a wide range of technology devices (Demers, Weiss-Lambrou & Ska, 1996).

The concept of satisfaction that guided the construction of QUEST is defined as a person's critical evaluation of several aspects of a device. This evaluation is

influenced by one's expectations, perceptions, attitudes and personal values. This definition is consistent with current satisfaction theoretical bases (Keith, 1998; Simon & Patrick, 1997) and was originally inspired by Linder-Pelz' (1982) theoretical work on satisfaction. In developing the instrument, data from several sources contributed to item generation including empirical research findings, clinical experience and a tridimensional framework of satisfaction derived from Scherer's Matching Person and Technology model (1996). Accordingly, the variables most likely to influence user satisfaction were related to the person, his or her environment, and specific characteristics of the technology device. Based on this theoretical premise as well as on content experts' feedback, the resulting QUEST instrument was comprised of 24 items related to particular characteristics of the device as well as to several psychosocial and service delivery attributes (Demers et al., 1996).

In concrete terms, QUEST is divided into three parts and is administered in a face-to-face interview. Part one consists of close-ended questions aimed at describing the context in which user assistive device satisfaction or dissatisfaction developed. In Part two, the evaluator asks the user to rate the degree of importance he or she attributes to those items (n=24) considered applicable to the assistive device being evaluated. As QUEST was designed to assess a wide range of assistive technology devices, not all items are applicable to every user and every situation. Importance is rated on a 5-point scale, with a score of 1 denoting "of no importance" and 5 indicating "very important". The user is also invited to add any other satisfaction items considered important. In Part three of QUEST, the person is asked to rate his or her satisfaction with 20 of the 24 items; four of the 24 items are not assessed for satisfaction because pre-

testing revealed that they were awkward to rate. Each item is scored with a 5-point satisfaction scale, with a score of 1 denoting “not satisfied at all” and 5 indicating “very satisfied”. For each item that is scored 3 (“more or less satisfied”), 2 (“not very satisfied”) or 1 (“not satisfied at all”), the user is asked to comment or explain the source(s) of dissatisfaction. Finally, the person is asked to score his or her overall satisfaction with the assistive technology device. The instrument was developed simultaneously in French and English, and a Dutch version translation was later constructed (Wessels, De Witte, Weiss-Lambrou, Demers, & Wijnhuizen, 1998).

In the course of the instrument’s ongoing development, some experimentation and field testing were conducted to assess its psychometric properties. With regards to reliability, individual items satisfaction ratings generated by QUEST Part three demonstrated moderate to substantial test-retest stability (weighted Kappa = 0.42 to 0.74) and fair to moderate interrater reproducibility (weighted Kappa = 0.30 to 0.66), over a period of 7 to 11 days. However, importance ratings derived from QUEST Part two revealed perceptibly weaker reliability coefficients (mean Kappa = 0.34) (Demers, Ska, Giroux & Weiss-Lambrou, submitted). With respect to validity, additional content validity evidence was obtained from the opinions and expertise of a group of assistive technology professionals in several countries (Demers, Wessels, Weiss-Lambrou, Ska, & De Witte, in press). A specifically constructed questionnaire was sent to 12 assistive technology experts who had tested and used QUEST in Canada, the United States and the Netherlands. Results of this international content validity study revealed that at the test level, QUEST was adequately sampled in terms of embracing all the important facets of satisfaction with assistive technology. At

the item level however, it was clear that minor changes needed to be made to the wording of some of the satisfaction variables and that the procedure used to administer QUEST needed to be modified.

In test construction, the eventual goal is to arrive at a test of minimum length that will yield scores with the desired degree of reliability and validity for the intended uses (Ferketich, 1991; Shrout & Yager, 1989). Typically, this is accomplished by field-testing a large pool of items and selecting a subset of items from the pool that makes the greatest contributions in regards to precise indicators (Crocker & Algina, 1986). In its present form, the QUEST tool has generated considerable interest from assistive technology providers and researchers. It was recently reviewed (RESNA, 1998; Smith, 1996) and used for data collection in a few Canadian studies (Quintin, Weiss-Lambrou & McCabe, 1997; Weiss-Lambrou, Demers, Tremblay, Ska, Lacoste, LeBlanc, & Dansereau, 1998), as well as in the Netherlands (Wessels et al., 1998) and Denmark (Brandt, 1998). Notwithstanding the relevance of the tool as a whole, its individual items may not be performing equally well and therefore measurement property discrepancies need to be closely examined. The purpose of the present study was to assess the psychometric properties of the 24 QUEST items and to examine if there is a subset of items that demonstrate optimal validity and reliability. For this purpose, the following methodological steps were conducted: identification of measurement properties to be assessed and the item parameters most relevant to these properties, data collection, estimation of the parameters, establishing a plan for the process of selecting items, selection of the final subset of items, and conducting a cross-validation study (Crocker & Algina, 1986). In the following section, the methods used to address these steps are described.

Methods

Measurement properties and related parameters

Ideally, a measurement instrument should consist of an optimal pooling or well-performing items. In the present study, six selective measurement properties were targeted. Three of the measurement properties were related to the validity of the individual items. The first criterion examined was **general acceptability** of the items, that is the extent to which they produces relevant responses. To determine if an item was considered acceptable to the device users, the number of non-applicable responses was analyzed. The second item property studied was **content validity** in terms of how important each of the items was to the assistive technology professionals and to the users of these devices. For this purpose, the degree of importance of each of the 24 items was rated by a group of content experts. Scoring was based on a three-point scale discriminating between “of primary importance”, “of secondary importance” and “neither”. The number of primary importance ratings scored by the experts was computed and used as an estimator of individual item relevance. Accordingly, the more substantial the agreement was among experts, the better was the item relevance. The ratings of item importance from the user viewpoint were also examined. In QUEST Part two, subjects were asked to rate the degree of importance attributed to those items considered applicable to the device being evaluated. Importance was rated on a 5-point scale, with a score of 1 denoting “of no importance” and 5 indicating “very important”. This parameter was obtained by averaging the importance scores obtained from the subjects. The third measurement property assessed, **criterion validity**, was based on correlational evidence with a criterion of interest. The parameter definition was however hampered by the lack of a “golden standard” measure of satisfaction

with assistive technology. A Spearman correlation coefficient between each item and the measure of overall satisfaction (which was in QUEST Part three) with device was computed as a gross estimate of the item's capacity to measure the concept under study as a positive relationship was expected.

The subsequent two measurement properties examined were concerned with reliability. Criterion four considered the item **contribution to internal consistency**. If an item has a sufficiently lower-than-average correlation with the other items, removing it raises the reliability of the scale as a whole (DeVellis, 1991). In the present study, this measurement property was investigated by examining item-total correlations, with the item of interest eliminated from the calculation of the total score. For the fifth measurement property, that is **test-retest stability**, the criterion used was Cohen's weighted Kappa (1968).

Finally, item **sensitivity** was set as the sixth measurement property of the item analysis process. Sensitivity was defined as the ability of the item to discriminate among individuals with varying levels of satisfaction (Ferketich, 1991). It was assessed in the amount of item dispersion in the data set, by comparing standard deviations statistics. In this context, larger values reflected more variability and consequently greater sensitivity.

Data sets

Data from three sources were used in the item selection process: a Montreal sample, a Dutch sample and data derived from an International Content Validation Study (Demers et al., in press).

The Montreal sample was first comprised of 158 subjects who were assistive technology users. The subjects were recruited through assistive technology centers in three Montreal rehabilitation facilities: *Centre de réadaptation Lucie-Bruneau*, *Institut de réadaptation de Montréal*, and *Centre de réadaptation Constance Lethbridge*. The assistive devices targeted were primarily seating and mobility aids, such as manual and powered wheelchairs and lower limb prostheses. QUEST was administered to the subjects by one of five trained interviewers on two separate occasions, with seven to eleven days between evaluation sessions. Data collected at the first administration of the instrument was used for this study as well as data from both sessions for the purpose of estimating test-retest stability. Not all the subjects were however evaluated twice by the same rater and consequently this parameter sample size was smaller. Before doing the item analysis, data were screened to detect any missing (whereby no subjects responded to less than 50% of the items) or aberrant data (whereby some subjects had zero variability in responses). Accordingly, the available sample was constituted of 150 subjects, including 85 subjects for test-retest stability.

The second source of data was a Dutch sample consisting of 375 device user subjects. The devices assessed included wheelchairs, scooters, shower seats, stairlifts, home adaptations and adapted beds. A large proportion of this sample (82%) was recruited from the TNO-PG Dutch Research Institute and the remaining 18% were selected from the Institute for Rehabilitation Research (IRV) in Hoensbroek, the Netherlands. All subjects were administered the Dutch version of QUEST (D-QUEST) which was equivalent to QUEST except for one aspect: the entire set of 24 items were rated for satisfaction instead of only 20

items as previously described. In both of these facilities, no formal training of the evaluators was provided and a large number of them (total of $n=31$) were involved in the data collection. Prior to conducting the analysis, the data set was inspected to ensure it was suitable for the intended purpose. Data were screened with regards to age of subjects (children were eliminated from sample), aberrant data (zero variability), missing data and non-applicable responses (individuals who responded to less than 50% of the 24 items were excluded). Based on these criteria, 32.5% of the respondents were eliminated and consequently the available Dutch sample was comprised of 253 subjects.

The third source of data was derived from the International Content Validation Study (Demers et al., in press). In this study, 12 international content experts were invited to participate in a multi-site testing and content validation of QUEST. These content experts represented 11 facilities in Canada, the United States and the Netherlands. They were each provided with a QUEST kit and were given six to eight months to use the satisfaction tool with any number of assistive devices available to them. Accordingly, QUEST was tested on a total of 578 assistive devices which were primarily seating and mobility aids (wheelchairs, scooters) together with transfer aids (toilet adaptation, shower seats and chairs, stair-lifts), augmentative communication devices, lower limb prostheses and environmental control devices. As content experts, the respondents completed a questionnaire on the instrument's content coverage and item relevance, the administration procedures, the instruction manual, the assessment materials and the future commercial publication of QUEST. Only the results obtained from the question on item relevance are presented in this paper.

Item selection plan

The item selection plan of the present study consisted of two parts. The first part served to define the decision criteria used for determining whether or not an individual item demonstrated an acceptable level of performance for each of the measurement properties identified. The second part involved decision making whereby these criteria were applied by establishing priorities for the measurement properties, and by introducing a factor analytic technique to support and finalize the selection of items.

Part one: Decision criteria

The decision criteria used to select the best performing QUEST items are summarized in Table I. These criteria were based on published standards and commonly used benchmarks, as well as on the researchers' judgements. The cutoffs that were used to sort and discriminate among the item pool are presented together with the related measurement property, computed parameter, and data source. As shown in this table, in order to assess general acceptability of an item, the number of non-applicable responses in the Montreal sample was analyzed for items 1 through 20, and in the Dutch sample for items 21 through 24. To define the cutoff criterion, the items were ranked in descending order of magnitude, from the most to the least acceptable. A clear division between large and small values was observed at approximately 10%. This percentage value was consequently chosen as the criterion to select items.

Content validity was studied from two viewpoints. The first perspective emanated from a group of assistive technology professionals and the relevant data was derived from the International Content Validation Study. A majority vote

Table I: Decision criteria for item selection

Measurement properties	Parameters	Sources of data	Decision criteria
1. General acceptability	Proportion of non applicable responses	Montreal sample (items 1 to 20) Dutch sample (items 21 to 24)	Percentage lower than 10%.
2. Content validity	Item relevance according to experts	International Content Validation Study	Agreement on primary importance of 50% and more.
	Item importance according to device users	Montreal sample	Average importance score of 4.00 or more.
3. Criterion validity	Correlation with overall satisfaction	Montreal sample (items 1 to 20) Dutch sample (items 21 to 24)	Correlation coefficient of 0.30 and higher.
4. Contribution to internal consistency	Item-total correlation	Montreal sample (items 1 to 20)	Correlation coefficient of 0.40 and higher.
5. Test-retest stability	Weighted Kappa coefficient	Montreal sample (items 1 to 20)	Coefficients of 0.50 and higher.
6. Sensitivity	Item dispersion	Montreal sample (items 1 to 20) Dutch sample (items 21 to 24)	Standard deviation of 1.00 and higher.

on the primary importance (50% and more experts) was required to retain a particular item. The other perspective originated from a group of users who had rated the importance attributed to the individual QUEST items on a five-point rating scale. Based on the word anchors provided with each response option, items whose averaged score did not obtain a score of at least 4.00 (indicating "quite important") were disqualified. In this case, the Montreal data set comprising 150 subjects was analyzed.

To assess criterion validity, the parameter chosen was a Spearman correlation coefficient between item and overall satisfaction. The analyses were conducted using the Montreal sample for items 1 through 20 and the Dutch sample for

items 21 to 24. Since the unidimensionality of satisfaction is not determined, the relevance of the criterion was not foolproof. In this context, significant ($p < 0.05$) values as low as 0.30 were judged sufficient for acceptance.

As previously stated, the item contribution to internal consistency was investigated by examining item-total correlations with the targeted item eliminated from the calculation of the total score. In the context of scale homogeneity, an item should correlate with the total score above 0.20 in order to reflect some covariance with the common factors among the items (Streiner & Norman, 1989). To introduce a stronger contrast between the QUEST items, those items that were attributed a coefficient of item-total correlations lower than 0.40 were eliminated. The correlation coefficients exploited the Montreal sample which means that they were computed for items 1 to 20.

The fifth measurement property was the test-retest stability coefficient. Judgements on the strength of agreement derived from weighted Kappa statistics were based on Landis and Koch' (1977) benchmarks for categorical data. Hence, coefficients values were interpreted according to the following categories: poor (<0.00), slight (0.00-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80) and almost perfect (0.81-1.00). In the present study, good performing items were expected to demonstrate moderate stability at 0.50 or a higher level of agreement. Test-retest stability was estimated for items 1 through 20 with the Montreal sample comprising 85 subjects as the source of data.

Finally, to assess item sensitivity, the range of the standard deviations derived from both the Montreal (items 1 to 20) and the Dutch (items 21 to 24) data sets

were compared. Just like that of general acceptability, the cutoff criterion was obtained after rank ordering the items from the most to the least sensitive item. There were no clear divisions observed in the standard deviations but a minimal value of 1.00 was defined as acceptable.

Part two: Decision making.

Item performance on the described measurement properties was used to eliminate the weakest items and retain the most appropriate subset for the revised QUEST. Although item performance on all six criteria were closely examined, item acceptability and content validity were considered as definite requisites. Good criterion validity was considered as an additional bonus however, if an item did not meet the cut-off point, it was not judged severely. It was expected that some items might have both good ratings and weaker ones, depending on the targeted measurement property. Decision-making for the inclusion or exclusion of the final item pool is not always easy to do solely by inspection of the data and consequently, further analysis of reliability and validity at the scale level needed to be conducted.

For this purpose, several factor analyses were performed with and without the questionable items to refine the content selection and explain the internal structure of satisfaction with assistive devices. Indeed, it was not certain that the resulting empirical structure would be congruent with the original tridimensional theoretical base of QUEST. Factor analysis is an analytical technique that permits the reduction of a certain number of interrelated variables to a smaller number of latent or hidden dimensions (Pedhazur & Schmelkin, 1991). Principal Axis Factoring (PAF), is the most widely used method of factor extraction for

explaining common variance (Pedhazur & Schmelkin, 1991) and it was used in this study. Because QUEST provided subjects with the option of scoring items as non-applicable, there was an important proportion of missing data. Therefore, in order to prevent a significant reduction of the sample size, a pairwise strategy was used to compute the matrix of inter-item correlation coefficients. To obtain a simple structure, items loading high on one factor and low on the remaining factors were needed. Meaningful factor loadings for each item were examined after both orthogonal (varimax) and oblique (oblimin) rotations. The results were similar, but orthogonal rotation was retained because it was easier to interpret. The reliability of the new QUEST scale was examined using the Cronbach coefficient alpha for internal consistency (Crocker & Algina, 1986). All of these analyses were based on the Montreal data set.

Cross-validation study

In order to test the appropriateness of the item pooling and the factorial model derived from the Montreal sample, a cross-validation analysis was carried out using a different sample of subjects. The Dutch data set (n=253) was used for this purpose and responses to the selected items were subjected to a replication factor analysis. Item assignments to a factor and loading values for both models were compared to evaluate the stability and the generalizability of the obtained internal structure solution. Factor extraction and correlation coefficients computations were performed using the same methodological features described earlier.

Results

Items performance

Results on the measurement properties of each of the QUEST items are presented in Table II. In terms of general acceptability, for 16 items (66,7%) the proportion of non-applicable responses was lower than 10% and for 12 items (50.0%), it was less than 5%. However, eight items were discarded by more than 15% of the subjects: *training* (#16; 29,3%), *device compatibility* (#17; 33,3%), *installation* (#19; 16.7%), *cost* (#20; 47,3%), *effort* (#21; 17.7%), *motivation* (#22; 16.5%), *reaction of others* (#23; 25.9%), and *social circle support* (#24; 29.2%). In this table's column "General Acceptability", items that were scored by more than 90% of subjects are indicated with a plus "+" sign while those that were untestable and frequently discarded (> 10%) are marked with a minus "-" sign. Italic font style reflects statistics obtained from the Dutch data set.

Item content validity was assessed separately according to content experts' perspective and device users' perspective. As displayed in Table II, degree of agreement on the primary importance of the items revealed that only two items were very weakly rated by a substantial proportion of content experts, namely *installation* (#19; 38%) and *reaction of others* (#23; 31%). The largest proportion of items (n=18, 75.0%) was considered of primary importance by half the respondents and more. These relevant items are marked with a plus "+" to distinguish them from the less relevant items which are identified by a minus "-" sign. From the users' point of view, analyses revealed that 19 of the 24 items generated a mean importance score above 4.00. Based on this cutoff, the items attributed with a minus "-" sign were: *training* (#16), *appearance* (#18), *cost* (#20), *reaction of others* (#23), and *social circle support* (#24).

Table II: Measurement properties of the 24 QUEST items

	General acceptability	Content validity (experts)	Content validity (users)	Criterion validity	Item-total correlation	Test-retest stability	Sensitivity
Item	% NA	% agreement	\bar{x}	Spearman r	Pearson r	WKappa	SD
1. Comfort	0.0% +	85% +	4.85 +	0.37** +	0.41 +	0.51 ** +	1.12 +
2. Dimensions	0.0% +	67% +	4.54 +	0.40** +	0.56 +	0.74 ** +	1.15 +
3. Professional service	1.3% +	67% +	4.58 +	0.37** +	0.51 +	0.61 ** +	1.21 +
4. Follow-up services	5.3% +	50% +	4.37 +	0.34** +	0.64 +	0.56 ** +	1.21 +
5. Simplicity of use	0.0% +	92% +	4.77 +	0.45** +	0.63 +	0.52 ** +	0.99 -
6. Effectiveness	0.0% +	92% +	4.83 +	0.53** +	0.63 +	0.62** +	0.84 -
7. Repairs/servicing	4.7% +	85% +	4.77 +	0.28** -	0.56 +	0.68 ** +	1.15 +
8. Durability	4.7% +	85% +	4.70 +	0.35** +	0.55 +	0.69** +	0.94 -
9. Adjustments	2.0% +	54% +	4.58 +	0.44** +	0.67 +	0.52 ** +	0.97 -
10. Safety	1.3% +	92% +	4.71 +	0.26** -	0.55 +	0.58 ** +	0.94 -
11. Service delivery	0.7% +	50% +	4.37 +	0.27** -	0.38 -	0.61 ** +	1.25 +
12. Weight	9.3% +	62% +	4.00 +	0.29** -	0.59 +	0.63 ** +	1.08 +
13. Multi-purpose fulness	4.0% +	46% -	4.41 +	0.46** +	0.41 +	0.58 ** +	1.15 +
14. Transportation	5.3% +	54% +	4.64 +	0.27** -	0.42 +	0.42 ** -	1.02 +
15. Maintenance	9.3% +	54% +	4.36 +	0.24 -	0.29 -	0.42** +	0.81 -
16. Training	29.3% -	69% +	3.76 -	0.25* -	0.40 +	0.62 ** +	1.37 +
17. Device compatibility	33.3% -	50% +	4.18 +	0.25* -	0.58 +	0.61 ** +	0.82 -
18. Appearance	1.3% +	46% -	3.75 -	0.26** -	0.45 +	0.48 ** -	0.82 -
19. Installation	16.7% -	38% -	4.40 +	0.25** -	0.49 +	0.49 ** -	0.86 -
20. Cost	47.3% -	46% -	3.10 -	0.21 -	0.17 -	0.55 ** +	0.98 -
21. Effort	17.7% -	75% +	4.42 +	0.34** +			1.06 +
22. Motivation	16.5% -	54% +	4.61 +	0.40** +			0.79 -
23. Reaction of others	25.9% -	31% -	2.84 -	0.12 -			0.91 -
24. Social circle support	29.2% -	46% -	3.87 -	0.03 -			0.66 -

Legend: **: $p < 0.01$; *: $p < 0.05$; NA: Non-applicable; \bar{x} : Mean; SD: Standard deviation

For the item criterion validity, positive Spearman's correlation coefficients (significant level of 0.01) were found for 18 of the 24 items. The strength of the associations were however rather weak, especially for the following items: *repairs/servicing* (#7), *safety* (#10), *service delivery* (#11), *weight* (#12), *transportation* (#14), *appearance* (#18), and *installation* (#19). Items were assigned a minus "-" sign when the correlation coefficients was either lower than 0.30 or non-significantly different from 0.00. As shown in this table, estimators derived from the Dutch data are presented in italic.

Item reliability results were firstly considered in terms of contribution to scale homogeneity, by way of item-total correlation. As shown in Table II, coefficients were available for 20 items and most of them were moderate and values ranged from 0.17 to 0.67. Items with the highest item-total correlations were: *follow-up services* (#4, 0.64), *simplicity of use* (#5, 0.63), *effectiveness* (#6, 0.63), and *adjustments* (#9, 0.67). The lowest item-total correlation was that found for *cost* (#20; 0.17). Because of listwise deletion of non-applicable data, the sample size available for this analysis was reduced to 55 subjects, when excluding the item *cost* (#20). The statistic for this item was based on a sample comprising only 20 subjects.

Estimated test-retest correlation coefficients ranged from 0.42 to 0.74. According to Landis and Koch' (1977), these results signify moderate to substantial agreement. Based on a more severe cutoff at 0.50, a minus "-" sign was assigned to *transportation* (#14; 0.42), *maintenance* (#15; 0.42), *appearance* (#18; 0.48) and *installation* (#19; 0.49). The items which achieved substantial agreement (0.61 to 0.80) were *dimensions* (#2), *professional services* (#3),

effectiveness (#6), repairs/servicing (#7), durability (#8), service delivery (#11), weight (#12), training (#16), and device compatibility (#17).

As illustrated in the far right-hand column of Table II, fairly homogeneous results were observed in terms of the sensitivity of the items. Extreme standard deviations (SD) values ranged from 0.66 for the weakest item (#24; *social circle support*) to 1.37 for the most sensitive item (#17 *training*). The plus “+” and minus “-” signs were used to highlight the item performance for this measurement property. Because they were derived from the Dutch data, the four last estimators appear in italic font style.

Selection of the best performing items

Upon inspection of Table II, it is clear that some items demonstrated more consistently desirable properties than others. These items (#1 to #11) are listed in the upper part of Table II. Although the item *safety* (#10) did not reach the cutoff set for criterion validity and sensitivity, it was selected because of its very high ratings on the remaining measurement properties. Similarly, *service delivery* (#11) did not prove to perform well on criterion validity and item-total correlation, however, these lower performances were outweighed by large values of acceptability and sensitivity. Another set of items (#15 to #24) demonstrated perceptibly weaker ratings with regards to some or all of the studied measurement properties. This second grouping of items are in the lower part of Table II. The three shaded remaining items (#12, #13, #14) evidenced both good ratings and weaker ones, depending on the targeted property. As anticipated, further analysis of reliability and validity at the scale level were conducted to help select the final item pool.

For this reason, several factor analyses were performed with and without the questionable items on the Montreal data set comprised of 150 subjects. This sample size exceeded the recommendation of having at least 10 times as many subjects as variables (Nunnally, 1978). Both the Bartlett Test of Sphericity ($p < 0.000$) and the KMO measure of sampling adequacy (approximately 0.81) demonstrated that the data were appropriate for the planned analysis. A two-factor structure consistently emerged, regardless of the number of variables that were subjected to the analysis (11 to 14). Slight variations in the amount of total item variance were explained (from 37.2% to 42%) and internal consistency of the scale and subscales (0.75 to 0.83) were however found. As more items were submitted to the analysis, two conflicting results were observed: the alpha tended to increase while the proportion of total variance explained tended to recede. Under these circumstances and after examining all conceivable combinations, a 12 item solution including the item *weight* (#12) was judged the best possible compromise.

The factor structure matrix shown in Table III, represents the loadings of the 12 items that comprise the final item selection with factors. The communalities, that is the proportion of variance that is accounted for by this solution are reported in the right-hand column. Small portions of *comfort* (#1), *safety* (#10), and *service delivery* (#11) variances were explained (respectively 0.193, and 0.224, and 0.180). For this analysis, a conservative threshold for meaningful loadings at 0.30 was employed (Pedhazur & Schmelkin, 1991). Results reveal that most items are high on one factor and low on the other, thus contributing positively to a simple resulting structure. Two items however performed slightly differently. Item *effectiveness* (#6) loaded on both factors but more substantially on Factor

one. Item *durability* (#8) loaded moderately on the two factors, somewhat more with Factor two. Because it is conceptually more closely related to the first cluster of items, *durability* was nevertheless assigned to Factor one. Based on the content of each scale, they were respectively named assistive technology DEVICE (consisting of eight items) and assistive technology SERVICES (consisting of four items). The DEVICE factor accounted for 20.6% of the explained common item variance and the SERVICES factor accounted for 18.7% for an explained common item variance totalizing 39.3%. Based on Principal Component Analysis (PCA), the total item variance explained by this solution attained 48.4%.

Table III: Results of factor analysis (after orthogonal rotation) of 12 satisfaction items, using the Montreal sample (n=150)

No	Item	DEVICE	SERVICES	Communalities
1.	Comfort	0.420		0.193
2.	Dimensions	0.608		0.381
5.	Simplicity of use	0.661		0.485
6.	Effectiveness	0.589	0.346	0.466
8.	Durability	0.361	0.419	0.306
9.	Adjustments	0.658		0.449
10.	Safety	0.396		0.224
12.	Weight	0.577		0.338
3.	Professional service		0.689	0.509
4.	Follow-up services		0.823	0.696
7.	Repairs/servicing		0.689	0.487
11.	Service delivery		0.394	0.180

The Cronbach' alpha coefficient of internal consistency reached 0.82 for the entire set of items, 0.80 for the DEVICE subscale and 0.76 for the SERVICES subscale. According to the acceptability levels proposed by DeVellis (1991), the first two coefficients are considered very good while the third estimator can be judged as being respectable. The mean item intercorrelations was 0.28.

Cross-validation

To conduct the cross-validation study, the 12 selected items were submitted to a replication factor analysis using a different sample of subjects. The Dutch sample comprised of 253 subjects was an appropriate sample size to conduct the proposed analysis (Nunally,1978). Both the Bartlett Test of Sphericity ($p < 0.000$) and the KMO measure of sampling adequacy (0,76) demonstrated that the data were appropriate for the planned analysis.

The replication study factor structure matrix presented in Table IV, represents the loadings of the 12 items with the factors. Once again, the communalities are reported in the right-hand column and a conservative threshold for meaningful loadings at 0.30 (Pedhazur & Schmelkin, 1991) was used. In comparison with the previous analysis, an even smaller proportion of *service delivery* (#11) variance was accounted for in this solution (0.108). Despite this small discrepancy, an identical two factor structure (DEVICE and SERVICES) emerged accounting for a very similar proportion of explained total common item variance (39,6%) and total item variance (49.0%). Three items showed high loadings on both factors: *effectiveness* (#6) and *weight* (#12), loaded more substantially on DEVICE while the item *durability* (#8), loaded slightly more with SERVICES but was again assigned to the DEVICE factor for conceptual

reasons. In comparison with the former analysis, Factor one accounted for a slightly more important share of the explained common item variance (24.6%), to the detriment of Factor 2 (15.1%).

The Cronbach' alpha coefficient of internal consistency was 0.84 for the whole set of items, 0.82 for the DEVICE subscale and 0.64 for the SERVICES subscale. The first two coefficients demonstrated a very good level of acceptability (DeVellis, 1991) while the latter one is perceptibly weaker.

Table IV: Results of factor analysis (after orthogonal rotation) of 12 satisfaction items, using the Dutch sample (n=253)

No	Item	DEVICE	SERVICES	Communalities
1.	Comfort	0.764		0.611
2.	Dimensions	0.611		0.399
5.	Simplicity of use	0.766		0.609
6.	Effectiveness	0.585	0.377	0.484
8.	Durability	0.339	0.395	0.271
9.	Adjustments	0.629		0.402
10.	Safety	0.467		0.305
12.	Weight	0.474	0.316	0.325
3.	Professional service		0.503	0.333
4.	Follow-up services		0.651	0.434
7.	Repairs/servicing		0.685	0.475
11.	Service delivery		0.312	0.108

Discussion

In the development of QUEST, close examination of its constitutive items was considered a logical and necessary step to establishing the validity and the reliability of the scale. Since several studies using QUEST had already been conducted, data from three sources were available to the researchers of this study. The methodology used examined item performances on six measurement properties, which is more than generally employed in item analysis (Crocker & Algina, 1986; Ferketich, 1991). Moreover, the selection of items was optimized by submitting potentially suitable subsets of items to factor analysis. By applying this dual approach, it is hoped that the selection of the best set of QUEST items was more robust than would normally be the case.

A meaningful and essential question to raise about a measure is whether it is consistent with the definition of the construct it is intended to be measuring. When studying satisfaction with assistive devices, it is necessary to examine whether the tridimensional conceptual framework of satisfaction was confirmed. Factor analysis is one of the most powerful analytical approaches that can be used for this purpose (Pedhazur & Schmelkin, 1991). The results of this study suggest that satisfaction with assistive technology is comprised of two unobserved latent dimensions, respectively named assistive technology DEVICE and SERVICES. The third hypothesized facet of the original conceptual scheme was not empirically supported. Indeed, personality-based items such as *motivation* and *social circle support*, among others, did not perform well in the item analysis. Such results do not suggest that sociocultural and motivational issues are not relevant, since very convincing evidence on their importance have been found in the literature (Galvin & Scherer, 1996; Kemp, 1993; Luborsky,

1993). What this study's findings do suggest however is that other suitable instruments need to be developed to specifically address these aspects of consumer evaluation of assistive technology. An example of such an instrument is the Psychosocial Impact of Assistive Devices (PIADS) (Day & Jutai, 1996) that addresses the psychosocial impacts of assistive technology on users' with regards to competence, adaptability, and self-esteem.

QUEST was found to be valid and reliable for assessing assistive technology device features and related service issues, as emphasized by the number of items that were selected and the significant loading values obtained. The DEVICE factor embraced eight items pertaining to salient characteristics of the assistive technology device: *comfort, dimensions, simplicity of use, effectiveness, durability, adjustments, safety, and weight*. From a theoretical viewpoint, the emergence of this factor supports and reinforces those works that view technology quality as a top priority in device selection (Cook & Hussey, 1995; Vanderheiden, 1987), use (Galvin & Scherer, 1996) and evaluation (Bain, 1997; Eblen, 1992, Kohn, Mortola & LeBlanc, 1991). The SERVICES factor encompassed four intercorrelated items: *professional services, follow-up services, repairs/servicing, and service delivery*. In the field of assistive technology, these variables are of paramount importance; they reflect the environmental influence on use (Galvin & Scherer, 1996) and they are assessed within the assistive technology system (Cook and Hussey, 1995).

In the literature, some concordant viewpoints are found with regards to the bidimensional structure of general satisfaction with assistive technology. In her systemic evaluation approach, Bain (1997) posited that assistive technology is comprised of devices and service delivery. In a keynote address to the American

Association of Occupational Therapy, Vanderheiden (1987) emphasized the proper choice of advanced technology and effective delivery as the essential conditions of successful assistive technology provision. Similarly, Kohn, Mortola and LeBlanc (1991) referred to two areas of practice: the provision of services, and the devices themselves. These authors suggested that satisfaction with either part of the system (devices and provision) does not imply satisfaction with the other, although they considered that if the assistive device is well received, it is likely that the service system will be perceived as satisfactory. In this context, one could expect that some items would have some loadings on both factors. In the present study, this was the case for the item *durability*, and to a lesser extent, the item *effectiveness*.

The fit of the proposed satisfaction model was judged as reasonably good. The communalities ranged from 0,180 to 0,696. Weaker values reflect low common item variance which can be explained by a certain degree of specificity in topics pertaining for example, to items such as *comfort*, *weight*, and *service delivery*. However, the two factor structure is not challenged by this phenomenon. The amount of total item variance explained by the two factors (48.4%) was very acceptable in light of other current satisfaction instruments (Keith, 1998; LeVois, Nguyen & Attkisson, 1981; Tanner, 1982; Williams, 1994). By revealing very similar results and an identical bidimensional structure, the cross-validation study gave additional strong support to the adequacy and stability of the satisfaction model.

In terms of internal consistency, the reliability of the scale (0,82 and 0.84) and subscales (0.64 to 0.82) can be considered positively. The Cronbach's alpha statistic is very much dependent upon the strength of the interitem correlations

and the number of items in a scale (Carmines & Zeller, 1979). For scales comprised of heterogeneous items, this statistic is also known to underestimate true reliability (Shrout & Yager, 1989). In the present study, the common variance was quite modest (mean interitem correlation of 0.28 for the Montreal sample), the number of items comprising the SERVICES subscale is small ($n=4$), and the scales comprised specific items. In this context, the results proved to be very good and they compared very favorably with other existing satisfaction instruments (Campen, Sixma, Friele, Kerssens, & Peters; 1995).

Most of the data used for the item selection came from the Montreal sample which focused on seating and mobility aids as well as on lower limb prostheses. Therefore, it can be postulated that the items might have been rated differently by users of other types of assistive devices. For example, the percentage of non-applicable responses to the item *weight* would certainly increase in the case of devices that are never or seldom moved, such as environmental control units or adapted car steering wheels. Similarly, the importance attributed by the users to individual items, which was one of the content validity parameters, is likely to vary according to type of device. Fortunately, the other parameters involved in this analysis were not so device-dependent. Therefore, the impact of this bias was likely to be of moderate or weak importance.

In conclusion, this study made an important contribution to the conceptual modeling of users' satisfaction with assistive technology. The findings revealed a stable two factor structure, related to assistive technology DEVICE and assistive technology SERVICES. In terms of measuring satisfaction along these dimensions, a challenge that needs yet to be faced is the sensitivity of the revised QUEST. In the present study, this measurement property was not

confirmed for the aggregated 12 items nor for the subscales. However, individual item variances were not considerable. It is hoped that in future studies, the 12 item QUEST resulting from this study's findings will discriminate between individuals with different satisfaction levels, over a large spectrum of disabilities and assistive devices.

Acknowledgments

This paper is based on the doctoral thesis of Louise Demers. The authors gratefully acknowledge the financial support provided by the *Fonds pour la recherche en santé du Québec*, the Canadian Occupational Therapy Foundation and the Université de Montréal. Special thanks are extended to Dr. Luc de Witte, Roelof Wessels, and the personnel of the Institute for Rehabilitation Research (IRV) and TNO-PG for sharing with us their Dutch data.

References

- Bain, B. K. (1997). Assistive technology in occupational therapy. In M. E. Neistadt & E. B. Crepeau (Eds.), *Williard and Spackman's occupational therapy* (9th ed., pp. 498-513). Philadelphia, PA: Lippincott.
- Brandt, A. (1998). Kunsten at male tilfredshed [Danish]. *HIT*, 2, 12-13.
- Campen, C., Sixma, H., Friele, R. D., Kerssens, J. J., & Peters, L. (1995). Quality of care and patient satisfaction: A review of measuring instruments. *Medical Care Research and Review*, 52, 109-133.
- Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. In J. L. Sullivan (Ed.), *Series in quantitative applications in the social sciences* (Number 07-017). Beverly Hills, CA: Sage.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement of partial credit. *Psychological Bulletin*, 70, 213-220.
- Cook, A. M., & Hussey, S. M. (1995). *Assistive technologies: Principles and practice*. St-Louis, MO: Mosby.
- Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. Orlando, FL: Harcourt Brace Javanovich.
- Day, H., & Jutai, J. (1996). Measuring the psychosocial impact of assistive devices: The PIADS. *Canadian Journal of Rehabilitation*, 9, 159-168.
- Demers, L., Ska, B., Giroux, F., & Weiss-Lambrou, R. (submitted). Stability and reproducibility of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Journal of Rehabilitation Outcomes Measurement*.
- Demers, L., Weiss-Lambrou, R., & Ska, B. (1996). Development of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Assistive Technology*, 8, 3-13.
- Demers, L., Wessels, R., Weiss-Lambrou, R., Ska, R., & De Witte, L. (in press). An international content validation of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Occupational Therapy International*, 6(3).

DeRuyter, F. (1995) Evaluating outcomes in assistive technology: Do we understand the commitment? *Assistive Technology*, 7, 3-16.

DeVellis, R. F. (1991). *Scale development - Theory and applications*. Newbury Park, CA: Sage.

Eblen, C. (1992). Issues involved in the evaluation of assistive devices. *Topics in Geriatric Rehabilitation*, 8(2), 6-11.

Ferketich, S. (1991). Focus on Psychometrics - Aspects of item analysis. *Research in Nursing and Health*, 14, 165-168.

Galvin, J. C., & Scherer, M. (1996). *Evaluating, selecting, and using appropriate assistive technology*. Gaithersburg, MD: Aspen.

Hall, J. A., & Dornan, M. C. (1988). Meta-analysis of satisfaction with medical care: Description of overall satisfaction levels. *Social Science and Medicine*, 27, 637-644.

Keith, R. A. (1998). Patient satisfaction and rehabilitation services. *Archives of Physical Medicine and Rehabilitation*, 79, 1122-1128.

Kemp, B. J. (1993). Motivational issues in the use of technology by older persons. *Technology and Disability*, 2, 65-70.

Kohn, J. G., Mortola, P., & LeBlanc, M. (1991). Clinical trials and quality control: Checkpoints in the provision of assistive technology. *Assistive Technology*, 3, 67-74.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.

LeVois, M., Nguyen, T. D., & Attkisson, C. C. (1981). Artifact in client satisfaction assessment - Experience in community mental health settings. *Evaluation and Program Planning*, 4, 139-150.

Linder-Pelz, S. (1982). Toward a theory of patient satisfaction. *Social Science and Medicine*, 16, 577-582.

Luborsky, M. R. (1993). Sociocultural factors shaping technology usage - Fulfilling the promise. *Technology and Disability*, 2, 71-78.

Nunally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Quintin, S., Weiss-Lambrou, R., & McCabe, D. (1997). *Enquête sur l'utilisation et la satisfaction concernant le lève-personne sur rail au plafond utilisé à domicile* [Survey of utilization and satisfaction with homebased lifts] (Research Report). Montreal, QC, Canada: Société Logique.

RESNA (Rehabilitation Engineering and Assistive Technology Society of North America). (1998). *RESNA resource guide for assistive technology outcomes: Assessment instruments, tools, & checklists from the field* (Vol. 2). Arlington, VA: RESNA.

Rubin, H. R. (1990). Patient evaluation of hospital care. *Medical Care*, 28, S3-S10.

Scherer, M. J. (1996). *Living in the state of stuck - How technologies affect the lives of people with disabilities* (2d ed.). Cambridge, MA: Brookline Books.

Shrout, P. E., & Yager, T. J. (1989). Reliability and validity of screening scales: Effect of reducing scale length. *Journal of Clinical Epidemiology*, 42, 69-78.

Simon, S. E., & Patrick, A. (1997). Understanding and assessing consumer satisfaction in rehabilitation. *Journal of Rehabilitation Outcomes Measurement*, 1(5), 1-14.

Smith, R. O. (1996). Measuring the outcomes of assistive technology: Challenge and innovation. *Assistive Technology*, 8, 71-81.

Streiner, D. L., & Norman, G. R. (1989). *Health measurement scales - A practical guide to their development and use*. Oxford, NY: Oxford University Press.

Tanner, B. A. (1982). A multi-dimensional client satisfaction instrument. *Evaluation and Program Planning*, 5, 161-167.

Trachtman, L. (1994). Outcome measures. Are we ready to answer the tough questions? (Editorial). *Assistive Technology*, 6, 91-92.

Vanderheiden, G. C. (1987). Service delivery mechanisms in rehabilitation technology. *American Journal of Occupational Therapy*, 41, 703-710.

Ware, J. E., Davies-Avery, A., & Stewart, A. L. (1978). The measurement and meaning of patient satisfaction. *Health and Medical Care Service Review*, 1, 1-15.

Weiss-Lambrou, R., Demers, L., Tremblay, C., Ska, B., Lacoste, & Dansereau, J. (1997). In QUEST of user satisfaction with assistive technology devices. In G. Anogianakis, G. Bülher, & M. Soede (Eds.), *Advancement of Assistive Technology* (pp. 428-431). Washington, DC: IOS Press.

Wessels, R. D., De Witte, L. P., Weiss-Lambrou, R., Demers, L., & Wijnhuizen, G. (1998). A Dutch version of QUEST (D-QUEST) applied as a routine follow-up within the service delivery process. In E. Placencia & E. Ballabio (Eds.), *Improving the Quality of Life for the European Citizen* (pp. 420-424). Washington, DC: IOS Press.

Williams, B. (1994). Patient satisfaction: A valid concept? *Social Sciences Medicine*, 38, 509-516.

CHAPITRE 5

Discussion générale et conclusions

L'objectif de la présente thèse était d'établir les valeurs métrologiques d'un instrument de mesure de la satisfaction envers les aides techniques utilisées en réadaptation, l'ÉSAT. Dans un premier temps, au chapitre 2, une étude d'envergure internationale sur la validité de contenu de l'instrument a confirmé la pertinence des éléments sélectionnés pour cerner le concept à l'étude de même que leur capacité à représenter correctement toutes ses facettes importantes. Par contre, comme il a été souligné, des améliorations visant à préciser le sens donné aux questions et à clarifier les procédures d'administration de l'ÉSAT sont souhaitées pour optimiser la validité des énoncés pris individuellement.

Afin de cerner les propriétés de mesure relatives à la fidélité de l'ÉSAT, la stabilité test-retest et la reproductibilité interjuges de chacun des énoncés ont été testées, analysées et discutées. Cette démarche, décrite dans le troisième chapitre, a révélé que les énoncés génèrent des résultats stables et reproductibles sur l'échelle ordinale comportant cinq niveaux de satisfaction mais qu'il en va autrement lorsque l'importance des énoncés, également recueillie au moyen d'une échelle ordinale à cinq niveaux, est considérée. Pour expliquer ces derniers résultats, la très faible dispersion des données autour des moyennes a été évoquée.

Finalement, au chapitre 4, une analyse approfondie des énoncés compris dans l'ÉSAT, portant sur leurs rendements individuel et collectif, a été réalisée. Les résultats ont permis de réduire de moitié le nombre des énoncés tout en optimisant les propriétés de mesure de fidélité et de validité de l'instrument pris dans son ensemble. Deux dimensions de satisfaction envers les aides techniques ont aussi été mises en évidence, c'est-à-dire la "technologie" et les "services".

Les résultats de chacune des études ont été discutés dans leurs chapitres respectifs et ils ne seront pas repris ici en détail. De plus, certains liens entre les résultats ont déjà été exposés dans le contexte du chapitre précédent, afin d'exploiter l'ensemble des informations pertinentes pour produire une analyse d'items qui soit la plus cohérente et significative possible. Par contre, il importe de replacer l'ensemble des résultats obtenus dans le contexte théorique général au sein duquel cette thèse s'inscrit, contexte qui a été présenté dans l'introduction. Les impacts des résultats rapportés précédemment seront donc à présent discutés en rapport avec les écrits sur les mesures des effets et impacts de l'utilisation d'aides techniques. En outre, la contribution que les résultats obtenus peuvent apporter au débat sur la conceptualisation de la satisfaction sera discutée.

1. L'évaluation des effets d'aides techniques

Études théoriques

La démarche de développement de l'ÉSAT coïncide avec l'apparition dans la littérature de quelques cadres conceptuels sur l'évaluation des effets et conséquences d'aides techniques. Tel que souligné en introduction, ces connaissances sont encore très fragmentaires et les nouvelles études réalisées sur ce thème ne pourront que raffiner les idées actuelles. Dans cette veine, les résultats des travaux de cette thèse supportent les modèles conceptuels existants en confirmant l'importance et la pertinence de s'attarder à l'évaluation de la satisfaction des usagers. Plus précisément, ils corroborent les propositions de Eblen (1992), Lawlor (1991) de même que Kohn et ses collaborateurs (1991) qui sont les premiers auteurs recensés à offrir des lignes directrices pour la mesure des effets et conséquences d'aides techniques. Ceux-ci s'entendent

pour inclure la satisfaction des usagers dans leurs structures évaluatives. Par conséquent, le fait que ce concept puisse être évalué de façon fidèle et valide renforce la présomption de validité de ces modèles.

Dans le même ordre d'idées, la nature de la présente recherche s'inscrit remarquablement bien dans le contexte des travaux d'auteurs plus récents. Comparativement à leurs prédécesseurs, DeRuyter (1995), Smith (1996b) de même que Scherer et Vitaliti (1997) définissent de façon encore plus précise la place que doit occuper la satisfaction d'usagers, parmi un ensemble d'autres catégories de mesures des effets et conséquences d'aides techniques. Smith (1996b) ainsi que Scherer et Vitaliti (1997) proposent notamment de considérer différents types d'utilisateurs des données évaluatives. Bien qu'il ne s'agisse pas d'un résultat de recherche escompté, il est intéressant de souligner que l'étude internationale de validation de contenu de l'ÉSAT (chapitre 2) confirme en partie la pertinence de ces types d'utilisateurs. En effet, des experts provenant de milieux cliniques, universitaires et industriels ont émis des avis et commentaires sur l'ÉSAT. Des points de vue divergeants ont été exprimés en regard de certaines variables et procédures propres à l'instrument. Cependant, la pertinence de la démarche évaluative dans son ensemble n'a jamais été remise en question par ces experts de provenance et de motivations diverses. En somme, les résultats de cette étude tendent à accréditer l'hypothèse de Smith (1996b) et de Scherer et Vitaliti (1997) selon laquelle l'évaluation de la satisfaction est utile à la fois pour les usagers, les chercheurs, les cliniciens et les administrateurs de programmes d'aides techniques.

Une incompatibilité entre l'approche économique et celle centrée sur les usagers semble se dessiner, à partir de certains résultats de notre démarche.

Comme il a été noté au premier chapitre, le modèle promulgué dans le cadre du projet "*CERTAIN*" (Oortwijn et al., 1997) s'intéresse au rapport coût/efficacité d'aides techniques sans aborder la perception de l'utilisateur ni son degré de satisfaction. Si l'approche économique ignore le point de vue de l'utilisateur, il en va cependant autrement des approches centrées sur le client, qui englobent les notions de coûts. En effet dans le cadre de leurs modèles respectifs, Smith (1996b), Scherer et Vitaliti (1997), Lawlor (1991) ciblent les coûts comme étant l'un des indicateurs susceptibles de refléter les effets et conséquences d'aides techniques. De même au sein de la version expérimentale de l'ÉSAT, se trouve un énoncé "coût" défini comme les frais occasionnés par l'achat, l'entretien et les réparations de l'aide technique. Au cours de la validation de contenu de l'instrument (chapitre 2), les résultats ont révélé que les experts ne s'entendent pas sur l'importance de cette variable. Quant au taux de réponses à cet énoncé (degré de satisfaction), il est très faible dans les échantillons de Montréal et des Pays-Bas, soit de moins de 50%. Ces résultats peuvent s'expliquer par des politiques locales de financement public des aides techniques, qui ne sont pas nécessairement les mêmes ailleurs au Canada, aux États-Unis ou aux Pays-Bas. Les coûts n'étant pas assumés par les usagers, il est concevable que ceux-ci y accordent moins d'importance. Dans ces conditions, le coût n'a pas été choisi comme faisant partie des items à retenir pour la prochaine version de l'ÉSAT (chapitre 4). En somme, les résultats convergent pour indiquer qu'il n'est pas approprié d'intégrer cette notion au sein d'une mesure de satisfaction. De ce fait, ils confirment une certaine incompatibilité entre l'approche économique et l'évaluation de la satisfaction.

Instruments de mesure

Dans l'introduction de ce travail, l'analyse de la littérature avait montré que même si les grandes catégories de mesures des effets et conséquences d'aides techniques apparaissent relativement bien circonscrites, les instruments de mesure eux-mêmes sont peu nombreux et ne répondent pas nécessairement bien aux besoins décelés dans le domaine. Entre autres, la pénurie d'instruments de mesure de la satisfaction de l'utilisateur avait été soulignée. La recension des écrits avait aussi permis de noter que les propriétés de mesure des quelques instruments retracés sont peu ou pas établies (Andrich et al., 1997, 1998; Day & Jutai, 1996; Smith, 1995), reflétant du même coup leur état de développement partiel. Dans la mesure où ils répondent à un besoin mis en évidence dès le départ, les résultats obtenus à travers la démarche méthodologique de l'ÉSAT sont définitivement progressifs. En effet, l'ÉSAT permettra d'obtenir des informations fiables étant données ses qualités métrologiques démontrées.

Cependant, même si les recherches reliées à cette thèse sont terminées, l'étude de l'ÉSAT devra se poursuivre. En effet, la validation d'un instrument de mesure est un processus perfectible et qui s'affirme à l'usage (Feldman et al., 1990; Messick, 1980). De plus, les propriétés de mesures de la version révisée de l'ÉSAT devront être précisées avec des données recueillies auprès de nouveaux échantillons de sujets. Il est cependant encourageant de noter que les stratégies utilisées et les résultats dévoilés aux chapitres précédents répondent à plusieurs des propriétés requises pour les instruments de mesure. Ainsi, la fidélité a été étudiée par le biais des trois approches définies par Contandriopoulos et ses collaborateurs (1990) c'est-à-dire, au chapitre 3, la stabilité test-retest et la

reproductibilité inter-juges et, au chapitre 4, la consistance interne. Les résultats se sont révélés très acceptables en ce qui a trait à l'échelle de satisfaction. En effet, les coefficients kappa pondérés ont oscillé entre 0,42 et 0,74 pour le test-retest et entre 0,30 à 0,66 pour la reproductibilité interjuges. L'homogénéité des items ($n=12$) a été estimée par la statistique alpha à 0,84. Quant à la validité, elle a été abordée en regard du contenu (chapitre 2) de même qu'en regard du construit de satisfaction (chapitre 4). La présomption de validité de l'instrument est soutenue par les résultats obtenus. Afin de poursuivre la validation de l'ÉSAT, des recherches futures pourront explorer la validité de critère de même que d'autres aspects de la validité de construit.

Par ailleurs, l'implantation progressive d'instruments de mesure nouveaux et variés, tels que le "PIADS" (Day & Jutai, 1996) pour cerner la qualité de vie, le "OT Fact" (Smith, 1995) pour préciser le niveau d'autonomie fonctionnelle, ou le "CERTAIN tool" (Andrich et al., 1997, 1998) pour rapporter la valeur coût/efficacité, permet de concevoir une batterie de tests complémentaires. Administrés conjointement avec l'ÉSAT, ces tests permettront de dresser un bilan exhaustif des effets et conséquences d'aides techniques, comme le suggèrent Scherer & Vitaliti (1997) de même que DeRuyter (1995). De plus, ils rendront possible l'élaboration et la vérification de diverses hypothèses de recherches sur les relations attendues entre les divers concepts ciblés. Par exemple, des relations positives entre la satisfaction, la qualité de vie et le degré d'autonomie fonctionnelle pourraient être examinées. Dans cette optique, les impacts des études réalisées pourront se répercuter sur les pratiques, via les cadres cliniques d'attribution des aides techniques (Bain, 1997; Cook & Hussey, 1995; Scherer, 1996a; Schoech et al., 1993), tout autant que sur le

développement des connaissances théoriques (DeRuyter, 1995; Eblen, 1992; Kohn et al., 1991; Lawlor, 1991; Scherer & Vitaliti, 1997; Smith, 1996b).

2. Le concept de satisfaction en réadaptation

En plus de contribuer au développement des connaissances sur les effets et conséquences d'aides techniques, la nature de l'étude réalisée au chapitre 4 permet d'aborder, pour la première fois, certaines questions propres à la structure du concept de satisfaction appliqué aux aides techniques utilisées en réadaptation.

Définitions

Les résultats obtenus au cours des chapitres précédents justifient un certain réaménagement des énoncés, procédures, et échelles d'évaluation pour la prochaine version de l'ÉSAT. Effectivement, des recommandations en ce sens ont été dégagées au cours de: (1) la validation de contenu de l'ÉSAT, visant à reformuler quelques énoncés, (2) les études de fidélité, éliminant l'évaluation de l'importance des énoncés, (3) l'analyse d'items, reconnaissant la faible sensibilité de certains énoncés et (4) les analyses factorielles, révélant une structure bidimensionnelle de la satisfaction.

À ce stade charnière des travaux, il est intéressant de questionner l'impact de la formulation des énoncés sur la définition des échelles de mesure de même que l'influence de celles-ci sur le concept mesuré (Pineault, communication personnelle, 10 mars 1999). Ainsi, la reformulation de certains énoncés peut amener un changement dans le type de réponse attendue. Par exemple, la question "Dans quelle mesure êtes-vous satisfait du service professionnel relié à

votre aide technique?” s’accompagne des options de réponses suivantes: pas satisfait du tout, peu satisfait, plus ou moins satisfait, assez satisfait et très satisfait. Cette question pourrait être réécrite comme un énoncé de la forme: “Les professionnels me donnent un service de qualité.” Dans ce cas, les options de réponses plausibles se définissent comme: pas du tout d’accord, plus ou moins d’accord, assez d’accord et entièrement d’accord. En adoptant la seconde formule, ne change-t-on pas du même coup la nature du concept mesuré? En effet, comment prétendre évaluer le degré de satisfaction lorsque ce mot n’apparaît plus dans l’énoncé présenté au sujet. Il semble plutôt se dégager une appréciation de la qualité. Pourtant, dans les deux alternatives, la même notion de services professionnels est ciblée. Il y a là matière à débat conceptuel, avant même d’aborder la problématique du point de vue méthodologique.

En effet, les mots employés pour définir les niveaux de l’échelle de mesure sont susceptibles d’influencer les distributions de réponses. Dans le chapitre d’introduction, la revue des études empiriques dénotait une propension pour de hauts taux de satisfaction (Bynum & Rogers, 1987; Finlayson & Havixbeck, 1992; Garber & Gregorio, 1989; Mann et al., 1993; Wilson et al., 1995; Winter & Keith, 1988). En réduisant le pouvoir de discriminer les sujets entre eux, cette situation crée un problème méthodologique important. Les résultats pourraient être différents si le terme “satisfaction”, à forte connotation positive, n’apparaissait pas systématiquement dans les choix de réponses. Voilà une autre question de recherche qui nécessite approfondissement. Pour y répondre, il s’agirait de comparer les résultats générés par des échelles concurrentes, pour des énoncés équivalents.

Ces questions conceptuelle et méthodologique interpellent un problème connu, celui du manque d'ancrage théorique de la satisfaction et l'ambiguïté persistante quant à sa signification profonde (Aharony & Strasser, 1993; Bond & Thomas, 1992; Carr-Hill, 1992; Keith, 1998; Pascoe, 1983; Williams, 1994). L'opérationnalisation d'un concept tel que la satisfaction passe par le choix minutieux des énoncés mais semble aussi dépendre, en grande partie, de l'échelle de mesure utilisée.

En somme, suite aux résultats de cette thèse, les mécanismes qui déterminent la satisfaction ne sont pas élucidés, loin de là. En tirant profit des conclusions dégagées du chapitre 4, il est possible de mieux comprendre la satisfaction envers les aides techniques utilisées en réadaptation. En effet, les résultats de l'analyse factorielle ont permis d'identifier une structure de satisfaction stable, comprenant deux dimensions distinctes, c'est-à-dire la "technologie" et les "services". Dès lors, par rapport aux aides techniques, il est possible de bonifier la définition de la satisfaction pour incorporer ces nouvelles notions. Dans le contexte de l'ÉSAT, celle-ci réfère désormais aux attitudes vis-à-vis la dimension "technologie" et la dimension "services" des aides techniques. Ces attitudes sont déterminées, dans une certaine mesure par les attentes, les perceptions, les standards personnels et les processus de comparaison individuels (Keith, 1998; Simon & Patrick, 1997).

Études théoriques

La découverte d'une structure bidimensionnelle de la satisfaction envers les aides techniques constitue un résultat majeur de cette thèse. Pour cette raison, il importe de replacer ce résultat dans le cadre des études théoriques de la

satisfaction, en mettant l'accent sur celles proposant une taxinomie de ses dimensions. En effet, à défaut de saisir précisément les mécanismes contribuant au développement de la satisfaction, plusieurs chercheurs se sont attardés à explorer empiriquement ses dimensions sous-jacentes (Baker, 1991; Carey & Seibert, 1993; Kristjanson, 1993; Shiloh, Avdor & Goodman, 1990; Tanner, 1982; Winter & Keith, 1988). Pour ce faire, la principale stratégie statistique utilisée est l'analyse factorielle, qui permet de regrouper ensemble les énoncés présentant une forte proportion de variance commune (Pedhazur & Schmelkin, 1991). Ce type d'analyse a été appliqué pour la première fois avec la satisfaction envers les aides techniques.

Comparativement aux autres mesures de la satisfaction développées dans le domaine de la santé, le phénomène étudié dans le cadre de cette thèse diffère par son originalité et sa spécificité. Par conséquent, les dimensions de mesure dégagées au chapitre 4, la "technologie" et les "services", sont différentes de celles relevées dans les champs de la médecine, de la gestion hospitalière ou des services infirmiers. Par exemple, les cinq dimensions de la pratique médicale identifiées par Baker (1991) inclut la continuité, l'accessibilité, la disponibilité, les soins médicaux et les lieux. Pour leur part, Shiloh et ses collaborateurs (1990) discernent des dimensions instrumentale, affective et procédurale à la satisfaction des clients de médecine génétique. Même si la nature conceptuelle des dimensions dévoilées distingue la présente étude des autres recherches, sa contribution à l'étude théorique de la satisfaction demeure significative pour le domaine de la réadaptation. En effet, elle renforce la pertinence d'une approche multidimensionnelle de la satisfaction. Ce faisant, les

résultats confirment également qu'il s'agit d'un principe directeur qui, dans le domaine de la santé, transcende les champs de spécialisation.

Cependant, un autre principe théorique qui semblait très important à l'origine du développement de l'ÉSAT n'a pas passé l'épreuve des tests empiriques. En effet, pour refléter le caractère arbitraire de la satisfaction, il avait été décidé de faire préciser l'importance attribuée à chacun des énoncés de satisfaction (Demers et al., 1996). Dans la version expérimentale de l'ÉSAT, cet objectif s'était concrétisé par la création de la partie 2, dans laquelle l'utilisateur devait indiquer le degré d'importance de chacun des énoncés sur une échelle de 1 à 5. Or, contrairement à ce qui était escompté, l'ensemble des données recueillies dans cette partie de l'évaluation indique qu'elle n'est ni fiable, ni valide. D'un point de vue théorique, ces résultats ne doivent pas être interprétés à l'effet que le caractère arbitraire de la satisfaction constitue une notion chimérique. En effet, d'autres chercheurs ont su démontrer de façon convaincante que le phénomène est très plausible (Carr-Hill, 1992; Pascoe, 1983; Vuori, 1991). Une meilleure explication repose sur les énoncés de satisfaction, qui ont été suffisamment bien sélectionnés pour être perçus comme importants par la grande majorité des usagers d'aides techniques. En conséquence de quoi la partie 2 de l'ÉSAT devient superflue.

Études empiriques

La réalisation de la démarche méthodologique présentée dans cette thèse n'est pas le fruit du hasard. Elle est née d'un mouvement d'intérêt pour la mesure de la satisfaction dans le domaine de la réadaptation (Heinemann et al., 1997; Keith, 1998; Simon & Patrick, 1997), décrit dans l'introduction de ce travail. Dans

le contexte de cette conclusion, il importe de souligner les impacts potentiels de nos résultats sur l'implantation d'études similaires dans le domaine.

La recension des écrits avait révélé que peu de données sur les propriétés de mesure des questionnaires de satisfaction sont susceptibles de sanctionner leur utilisation (Hammel, 1996; Keith, 1998). Or, les résultats des chapitres 2 et 3 de cette thèse fournissent aux utilisateurs de l'ÉSAT les informations nécessaires pour juger de la pertinence de l'instrument par rapport à leurs propres objectifs d'évaluation. Dans cette veine, la démarche suivie, inspirée à l'origine du cadre méthodologique en quatre phases de Benson et Clark (1982) est compatible avec les recommandations de chercheurs ayant publié récemment (Dobrzykowski, 1997; Keith, 1998; Osborn, 1998; Simon & Patrick, 1997). Ceux-ci proposent des balises pour encadrer la démarche de développement d'instruments de mesure. Par exemple, Dobrzykowski (1997) élabore sur les méthodes à suivre pour mesurer les résultats des interventions de réadaptation, Simon et Patrick (1997) décrivent une méthodologie de développement d'instruments de mesure de la satisfaction et Keith (1998) souligne les caractéristiques, embûches et points forts de l'évaluation de la satisfaction en réadaptation. Ces auteurs considèrent que la rigueur méthodologique constitue une condition importante pour l'avancement des connaissances en réadaptation, ce qui est congruent avec la réalisation de cette recherche.

Les résultats de cette thèse amènent à juger avec circonspection les résultats générés par plusieurs études sur la satisfaction dans le domaine de la réadaptation. En raison des lacunes inhérentes à plusieurs des instruments de mesure utilisés, les interprétations qui s'ensuivent sont empreintes d'incertitude. Par exemple, il est hasardeux de conduire des tests sur des scores totaux,

dérivés d'énoncés pour lesquels la structure unidimensionnelle n'a pas été vérifiée. Les énoncés doivent nécessairement être corrélés pour définir un concept sous-jacent commun. Il est souhaité que les travaux réalisés encouragent d'autres initiatives de développement d'instruments de mesure. Dans l'avenir, il est essentiel que le domaine de la réadaptation dispose de meilleurs outils de satisfaction pour faire la démonstration des retombées positives des interventions rendues.

3. Éléments de réponse et nouvelles questions

La contribution de cette thèse s'apprécie relativement à plusieurs aspects. Quelques uns ont été soulignés dans chacun des articles, alors que d'autres ont été plus clairement identifiés dans ce chapitre. Pour résumer, ce travail a permis de:

- Définir les champs de connaissances pertinents au développement d'un instrument de mesure de la satisfaction d'aides techniques, incluant l'évaluation des effets et conséquences de l'utilisation d'aides techniques, les notions reliées à la satisfaction des usagers de même que la métrologie;
- Établir la validité de contenu de l'ÉSAT dans une perspective d'application transculturelle et par divers types d'utilisateurs. Cette validation a porté sur les énoncés composant l'instrument de même que sur les procédures utilisées;
- Établir la fidélité test-retest et la reproductibilité interjuges des énoncés compris dans l'ÉSAT, permettant aux utilisateurs de la version expérimentale de juger de sa pertinence pour la réalisation de leurs propres études;
- Générer les informations nécessaires à la construction d'une version révisée de l'ÉSAT, ayant des propriétés de mesure optimales. Cette deuxième génération de l'instrument sera plus concise et pourra être auto-administrée,

ce qui ouvre la voie à des études épidémiologiques très variées;

- Contribuer à la compréhension du phénomène de satisfaction des aides techniques en mettant en lumière une structure stable composée de deux dimensions distinctes, la “technologie” et les “services”;
- Confirmer le modèle multidimensionnel de la satisfaction prôné dans le domaine de la santé;
- Appuyer les modèles théoriques de l'évaluation des effets et conséquences de l'utilisation d'aides techniques dans le champ de la réadaptation. Cette contribution passe par l'affirmation de la prime importance de la satisfaction;
- Contribuer au développement d'un corpus de connaissances empiriques et théoriques sur les aides techniques, en proposant une façon de mesurer leurs impacts;
- Répondre à une demande exprimée aux plans national et international pour une mesure standardisée de la satisfaction envers les aides techniques.

Par ailleurs, comme tout travail de recherche, la présente thèse suscite autant de questions qu'elle apporte de réponses. Par exemple, au terme de ces travaux, il demeure important de:

- Construire et valider la nouvelle version de l'ÉSAT, dans les langues des pays qui souhaitent l'utiliser;
- Vérifier la sensibilité de l'ÉSAT à détecter des changements dans le degré de satisfaction;
- Utiliser d'autres stratégies de validation de l'ÉSAT, tel que la validation de critère et la validation nomologique;
- Préciser le concept de la satisfaction envers les aides techniques du point de

vue de ses mécanismes et déterminants;

- Approfondir les distinctions perçues entre l'évaluation de la satisfaction et l'évaluation de la qualité en comparant, par exemple, les résultats obtenus sur des échelles concurrentes;
- Échafauder et tester des hypothèses de recherches sur les relations possibles entre la satisfaction, la qualité de vie, l'autonomie fonctionnelle ainsi que d'autres concepts liés à l'efficacité d'aides techniques.

Ces thèmes pourront faire l'objet de travaux futurs, permettant ainsi de poursuivre la démarche dont la maîtrise et la thèse sont le point de départ. Les travaux visant à construire la version révisée de l'ÉSAT seront menés en priorité.

En conclusion, cette thèse tient sa valeur du fait que l'ÉSAT permettra d'exercer une meilleure évaluation des effets et conséquences d'aides techniques attribuées aux personnes présentant une déficience, une incapacité ou une situation de handicap. Des retombées positives sont anticipées relativement à l'évaluation, l'attribution, le suivi, la recherche et le développement d'aides techniques répondant aux besoins des usagers.

Références

- Aharony, L., & Strasser, D. (1993). Patient satisfaction: What we know and what we still need to explore. *Medical Care Review*, *50*, 49-79.
- Andrich, R., Ferrario, M., & Moi, M. (1997). Cost-outcomes analysis in the provision of assistive technology to individual cases. Dans G. Anagnonakis, C. Bühler, & M. Soede (Eds.), *Advancement of assistive technology* (pp. 386-390). Amsterdam, The Netherlands: IOS Press.
- Andrich, R., Ferrario, M., & Moi, M. (1998). A model of cost-outcomes analysis for assistive technology. *Disability and Rehabilitation*, *20*(1), 1-24.
- Armstrong, G. D. (1981). The intraclass correlation as a measure of interrater reliability of subjective judgments. *Nursing Research*, *30*, 314-315, 320A.
- Bain, B. K. (1997). Assistive technology in occupational therapy. Dans M. E. Neistadt & E. B. Crepeau (Eds.), *Williard and Spackman's occupational therapy* (9th ed., pp. 498-513). Philadelphia, PA: Lippincott.
- Baker, R. (1991). The reliability and criterion validity of a measure of patients' satisfaction with their general practice. *Family Practice*, *8*, 171-177.
- Batavia, A. I., DeJong, G., & McKnew, L. B. (1991). Toward a national personal assistance program: The independent living model of long term care for persons with disabilities. *Journal of Health Politics, Policy and Law*, *16*, 523-545.
- Batavia, A. I., & Hammer, G. S. (1990). Toward the development of consumer-based criteria for the evaluation of assistive devices. *Journal of Rehabilitation Research and Development*, *7*, 425-436.
- Benson, J., & Clark, F. (1982). A guide for instrument development and validation. *American Journal of Occupational Therapy*, *36*, 789-800.
- Bilodeau, S. (1994). *L'utilisation de la prothèse chez la personne âgée amputée d'un membre inférieur*. Mémoire de maîtrise non publié, Faculté des lettres et sciences humaines, Université de Sherbrooke, QC.

Bilodeau, S., Hébert, R., & Desrosiers, J. (1998). Questionnaire sur la satisfaction des personnes amputées du membre inférieur face à leur prothèse: Développement et validation. *Revue Canadienne d'Ergothérapie*, 66, 23-31.

Bond, S., & Thomas L.H. (1992). Measuring patients' satisfaction with nursing care. *Journal of Advanced Nursing*, 17, 52-63.

Braekhus, A., Laake, K., & Engedal, K. (1992). The mini-mental state examination: Identifying the most efficient variables for detecting cognitive impairment in the elderly. *Journal of American Geriatrics Society*, 40, 1139-1145.

Bramadat, I. J., & Driedger, M. (1993). Satisfaction with childbirth: Theories and methods of measurement. *Birth*, 20, 2229.

Brandt, A. (1998). Kunsten at male tilfredshed [Danish]. *HIT*, 2, 12-13.

Brosseau, L., Philippe, P., Dutil, E., & Boulanger, Y. L. (1996). Mesure d'Indépendance Fonctionnelle (M.I.F.): Recension des écrits. *Journal de Réadaptation Médicale*, 16(1), 9-21.

Bynum, H., & Rogers, J. (1987). The use and effectiveness of assistive devices possessed by patients seen in home care. *Occupational Therapy Journal of Research*, 7, 181-191.

Canadian Association of Occupational Therapy. (1997). *Enabling occupation - An occupational therapy perspective*. Ottawa, ON: CAOT Publications ACE.

Canadian Burden of Illness Study Group. (1998). Burden of illness of multiple sclerosis: Part 2: Quality of life. *Canadian Journal of Neurologic Sciences*, 25, 31-38.

Campen, C., Sixma, H., Friele, R. D., Kerssens, J.J., & Peters, L. (1995). Quality of care and patient satisfaction: A review of measuring instruments. *Medical Care Research and Review*, 52, 109-133.

Carey, R. G., & Seibert, J. H. (1993). A patient survey system to measure quality improvement: Questionnaire reliability and validity. *Medical Care*, 31, 834-845.

- Carmine, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. Dans J. L. Sullivan (Ed.), *Series in quantitative applications in the social sciences* (Number 07-017). Beverly Hills, CA: Sage.
- Carr-Hill, R. A. (1992). The measurement of patient satisfaction. *Journal of Public Health Medicine, 14*, 236-249.
- Chamberlain, M. A., & Gallop, J. (1988). The disabled living center: What does it do? *British Medical Journal, 297*, 1523-1526.
- Chang, L. H., & Hasselkus, B. R. (1997). Occupational therapists' expectations in rehabilitation following stroke: Sources of satisfaction and dissatisfaction. *American Journal of Occupational Therapy, 52*, 629-637.
- Cicchetti, D. V. (1976). Assessing inter-rater reliability for rating scales: Resolving some basic issues. *British Journal of Psychiatry, 129*, 452-456.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*, 213-220.
- Conseil consultatif sur les aides technologiques. (1994). *Attribution d'aides techniques aux personnes âgées: État de la situation*. Québec: Auteur.
- Contandriopoulos, A. P., Champagne, F., Potvin, L., Denis, J. L., & Boyle, P. (1990). *Savoir préparer une recherche, la définir, la structurer, la financer*. Montréal, QC: Les Presses de l'Université de Montréal.
- Cook, A. M., & Hussey, S. M. (1995). *Assistive technologies: Principles and practice*. St-Louis, MO: Mosby.
- Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York, NY: Holt, Rinehart & Winston.
- Daly, R., & Flynn, R. J. (1985). A brief consumer satisfaction scale for use in inpatient rehabilitation programs. *International Journal of Rehabilitation Research, 8*, 335-338.

- Datta, D., & Ariyaratnam, R. (1996). Activities and users' views of a special seating clinic. *Disability and Rehabilitation*, 18, 365-368.
- Davies, D., & Hobbs, G. (1989). Measuring outpatient satisfaction with rehabilitation services. *Quality Review Bulletin*, 15, 192-197.
- Day, H., & Jankey, S. G. (1996). Lessons from the literature: Towards a holistic model of quality of life. Dans R. Renwick, I. Brown, & M. Nagler (Eds.), *Quality of life in health promotion and rehabilitation: Conceptual approaches, issues and applications* (pp. 39-50). Thousand Oaks, CA: Sage.
- Day, H., & Jutai, J. (1996). Measuring the Psychosocial Impact of Assistive Devices: The PIADS. *Canadian Journal of Rehabilitation*, 9, 159-168.
- Demers, L. (1995). Le développement d'une évaluation de la satisfaction des personnes âgées envers leurs aides techniques. Mémoire de maîtrise non publié, Université de Montréal, Montréal.
- Demers, L., & Giroux, F. (1997). Validité de la Mesure de l'Indépendance fonctionnelle (MIF) pour les personnes âgées suivies en réadaptation. *Revue Canadienne du Vieillissement*, 16, 626-646.
- Demers, L., Ska, B., Giroux, F., & Weiss-Lambrou, R. (submitted). Stability and reproducibility of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Journal of Rehabilitation Outcomes Measurement*.
- Demers, L., Weiss-Lambrou, R., & Ska, B. (1996). Development of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Assistive Technology*, 8, 3-13.
- Demers, L., Wessels, R., Weiss-Lambrou, R., Ska, R., & De Witte, L. (in press). An international content validation of the Quebec user evaluation of satisfaction with assistive technology (QUEST). *Occupational Therapy International*.
- DeRuyter, F. (1995) Evaluating outcomes in assistive technology: Do we understand the commitment? *Assistive Technology*, 7, 3-16.
- DeVellis, R. F. (1991). *Scale development - Theory and applications*. Newbury Park, Ca: Sage.

- Dobrzykowski, E. A. (1997). The methodology of outcomes measurement. *Journal of Rehabilitation Outcomes Measurement*, 1(1), 8-17.
- Donabedian, A. (1980). *Explorations in quality assessment and monitoring. Volume 1. The definition and approaches to its assessment*. Ann Arbor, MI: Health Administration Press.
- Eblen, C. (1992). Issues involved in the evaluation of assistive devices. *Topics in Geriatric Rehabilitation*, 8(2), 6-11.
- Elbeck M. (1987). An approach to client satisfaction measurement as an attribute of health service quality. *Health Care Management Review*, 12(3), 47-52.
- Elliott-Burke, T., & Pothast, L. (1997a). Measuring patient satisfaction in outpatient orthopedic setting, part 1: Key drivers and results. *Journal of Rehabilitation Outcomes Measurement*, 1(1), 18-25.
- Elliott-Burke, T., & Pothast, L. (1997b). Measuring patient satisfaction in outpatient orthopedic setting, part 2: Utilizing data to improve quality and market services. *Journal of Rehabilitation Outcomes Measurement*, 1(2), 16-22.
- Feeney, R. J. (1987). Designing for disabled persons. *International Journal of Disability Studies*, 9, 92-95.
- Feinstein, A. R. (1987). *Clinimetrics*. New Haven, CT: Yale University Press.
- Feldman, A. B., Haley, S., & Coryell, J. (1990). Concurrent and construct validity of the pediatric evaluation disability inventory. *Physical Therapy*, 70, 602-610.
- Ferketich, S. (1991). Focus on psychometrics - Aspects of item analysis. *Research in Nursing & Health*, 14, 165-168.
- Finlayson, M., & Havixbeck, K. (1992). A post-discharge study on the use of assistive devices. *Canadian Journal of Occupational Therapy*, 59, 201-207.
- Fisher, A. G., Liu, Y., Velozo, C. A., & Pan, A. W. (1992). Cross-cultural assessment of process skills. *American Journal of Occupational Therapy*, 46, 876-885.

Fitzpatrick, R. (1991). Surveys of patient satisfaction: I - Important general considerations. *British Medical Journal*, 302, 887-889.

Fitzpatrick R, & Hopkins A. (1983). Problems in the conceptual framework of patient satisfaction research: An empirical exploration. *Sociology of Health and Illness*, 5, 297-311.

Fleiss, J. L., Cohen, J., & Everitt, B. S. (1969). Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 72, 323-327.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatry Research*, 12, 189-198.

Fuhrer, M. J., Rintala, D. H., Hart, K. A., Clearman, R., & Young, M. E. (1992). Relationship of life satisfaction to impairment, disability, and handicap among persons with spinal cord injury living in the community. *Archives of Physical Medicine and Rehabilitation*, 73, 552-557.

Galvin, J. C., & Scherer, M. (1996). *Evaluating, selecting, and using appropriate assistive technology*. Gaithersburg, MD: Aspen.

Garber, S. L., & Gregorio, T. L. (1989). Upper extremity assistive devices: Spinal cord-injured patients with quadriplegia. *American Journal of Occupational Therapy*, 44, 126-31.

Gompertz, P., Pound, P., & Ebrahim, S. (1994). Validity of the extended activities of daily living scale. *Clinical Rehabilitation*, 8, 275-280.

Guillemin, B. B., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology*, 46, 1417-1432.

Guion, R. M. (1977). Content validity - The source of my discontent. *Applied Psychological Measurement*, 1(1), 1-10.

Hachisuka, K., Dozona, H., Ohmine, S., Shitama, H., & Shinkuda, K. (1998). Total surface bearing below-knee prosthesis: Advantages, disadvantages, and clinical implications. *Archives of Physical Medicine and Rehabilitation*, 79, 783-789.

- Haddad, S. (1991). Évaluation de la fiabilité et de validité d'un index d'humanisation des soins. Mémoire de maîtrise non publié, Département de médecine sociale et préventive, Université de Montréal, QC.
- Hall, J. A., & Dornan, M. C. (1988). Meta-analysis of satisfaction with medical care: Description of overall satisfaction levels. *Social Science and Medicine*, 27, 637-644.
- Hall, J. A., Milburn, M. A., & Epstein, A. M. (1993). A causal model of health status and satisfaction with medical care. *Medical Care*, 31, 84-94.
- Hammel, J. (1996). What's the outcome? Multiple variables complicate the measurement of assistive technology outcomes. *REHAB Management*, 9(2), 97-99.
- Harvey, R. F., Silverstein, B., Venzon, M. A., Kilgore, K. M., Fisher, W. P., Steiner, M., & Harley, J. P. (1992). Applying psychometric criteria to functional assessment in medical rehabilitation: III. Construct validity and predicting level of care. *Archives of Physical Medicine and Rehabilitation*, 73, 887-892.
- Hazard, R. G., Haugh, L. D., Green, P. A., & Jones, P. L. (1994). Chronic low back pain - The relationship between patient satisfaction and pain, impairment, and disability outcomes. *Spine*, 19, 881-887.
- Heinemann, A. W., Bode, R., Cichowski, K. C., & Kan, E. (1997). Measuring patient satisfaction with medical rehabilitation, *Journal of Rehabilitation Outcomes Measurement*, 1(4), 52-65.
- Hsieh, M. O., & Kagle, J. D. (1991). Understanding patient satisfaction and dissatisfaction with health care. *Health and Social Work*, 16, 281-290.
- Jacques, G. E., Ryan, S., Naumann, S., Milner, M., & Cleghorn, W. L. (1994). Application of quality function deployment in rehabilitation engineering. *IEEE Trans Rehabilitation Engineering*, 2, 158-64.
- Keith, R. A. (1995). Conceptual basis of outcome measures. *American Journal of Physical Medicine and Rehabilitation*, 74, 73-80.
- Keith, R.A. (1998). Patient satisfaction and rehabilitation services. *Archives of Physical Medicine and Rehabilitation*, 79, 1122-1128.

- Kemp, B. J. (1993). Motivational issues in the use of technology by older persons. *Technology and Disability, 2*, 65-70.
- Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health and disease. *Journal of Chronic Disease, 38*, 27-36.
- Koch, L. C., & Merz, M. A. (1995). Assessing client satisfaction in vocational rehabilitation program evaluation: A review of instrumentation. *Journal of Rehabilitation, 61*, 24-30.
- Kohn, J., Enders, S., Preston, J., & Motloch, W. (1983). Provision of assistive equipment for handicapped persons. *Archives of Physical Medicine and Rehabilitation, 64*, 378-381.
- Kohn, J. G., LeBlanc, M., & Mortola, P. (1994). Measuring quality and performance of assistive technology: Results of a prospective monitoring program. *Assistive Technology, 6*, 120-125.
- Kohn, J. G., Mortola, P., & LeBlanc, M. (1991). Clinical trials and quality control: Checkpoints in the provision of assistive technology. *Assistive Technology, 3*, 67-74.
- Korpela, R., Seppänen, R. L., & Koivikko, M. (1993). Rehabilitation service evaluation: A follow-up of the extent of use of technical aids for disabled children. *Disability and Rehabilitation, 15*, 143-150.
- Kramer, A. M. (1997). Rehabilitation care and outcomes from the patient's perspective. *Medical Care, 35*, js48-js57, Supplement.
- Kristjanson, L.J. (1993). Validity and reliability testing of the FAMCARE scale: Measuring family satisfaction with advanced cancer care. *Social Science Medicine, 36*, 693-701.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.
- Law, M. (1987). Measurement in occupational therapy: Scientific criteria for evaluation. *Canadian Journal of Occupational Therapy, 54*, 133-138.

- Lawlor, M. (1991). Evaluation of the impact of technology and innovative products in the clinical setting. *Occupational Therapy Practice*, 2, 74-86.
- Liang, M. H., & Jette, A. M. (1981). Measuring functional ability in chronic arthritis. *Arthritis Rheumatism*, 24, 80-86.
- LeVois, M., Nguyen, T. D., & Attkisson, C. C. (1981). Artifact in client satisfaction assessment - Experience in community mental health settings. *Evaluation and Program Planning*, 4, 139-150.
- Linder-Pelz, S. (1982a). Toward a theory of patient satisfaction. *Social Science and Medicine*, 16, 577-582.
- Linder-Pelz, S. (1982b). Social psychological determinants of patient satisfaction: A test of five hypotheses. *Social Science and Medicine*, 16, 583-589.
- Locker, D., & Dunt, D. (1978). Theoretical and methodological issues in sociological studies of consumer satisfaction with medical care. *Social Science and Medicine*, 12, 283-292.
- Lorentsen, O. (1997). Goal attainment - A tool to assess outcomes of assistive devices. Dans G. Anogianakis, C. Bühler, & M. Soede (Eds.) *Advancement of Assistive Technology* (pp. 102-104). Amsterdam, The Netherlands: IOS Press.
- Luborsky, M. R. (1993). Sociocultural factors shaping technology usage - Fulfilling the promise. *Technology and Disability*, 2, 71-78.
- Maclure, M., & Willett, W. (1987). Misinterpretation and misuse of the kappa statistic. *American Journal of Epidemiology*, 126, 161-169.
- Mann, W. C., Hurren, D., & Tomita, M. (1993). Comparison of assistive device use and needs of home-based older persons with different impairments. *American Journal of Occupational Therapy*, 47, 980-987.
- McComas, J., Kosseim, M., & Macintosh, D. (1995). Client-centered approach to develop a seating clinic satisfaction questionnaire: A qualitative study. *American Journal of Occupational Therapy*, 49, 980-985.
- McDaniel, C., & Nash, J. G. (1990). Compendium of instruments measuring patient satisfaction with nursing care. *Quality Review Bulletin*, 16, 182-188.

- McGrath, P. J., Goodman, J. T., Cunningham, S. J., MacDonald, B. J., Nichols, T. A., & Unruh, A. (1985). Assistive devices: Utilization by children. *Archives of Physical Medicine and Rehabilitation*, 66, 430-432.
- McMillan, J. R. (1987). Measuring consumer satisfaction to improve quality of care. *Health Progress*, 68(March), 54-55, 76, 78, 80.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Minkel, J. L. (1996). Assistive technology and outcome measurement: Where do we begin? *Technology and Disability*, 5, 285-288.
- Nunally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Office des personnes handicapées du Québec. (1993). *Le Québec: Porte ouverte de la communauté économique européenne en Amérique*. (Document de travail DPER-920). Québec: auteur.
- Oortwijn, W., Boxsel, J. V., & Beekun, T. V. (1997). Technology assessment and assistive technology. Dans G. Anagionakis, C. Bühler, & M. Soede (Eds.), *Advancement of assistive technology* (pp. 98-101). Amsterdam, The Netherlands: IOS Press.
- Osborn, C. E. (1998). Developing instruments for assessment of patient outcomes. *Journal of Rehabilitation Outcomes Measurement*, 2(6), 18-25.
- Pascoe, G. C. (1983). Patient satisfaction in primary health care: A literature review and analysis. *Evaluation and Program Planning*, 6, 185-210.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.
- Phillips, B., & Zhao, H. (1993). Predictors of assistive technology abandonment. *Assistive Technology*, 5, 36-45.
- Post, K. M. (1993). The promise of assistive technology. *American Journal of Occupational Therapy*, 47, 965-967.

Quintin, S., Weiss-Lambrou, R., & McCabe, D. (1997). Enquête sur l'utilisation et la satisfaction concernant le lève-personne sur rail au plafond utilisé à domicile [Survey of utilization and satisfaction with homebased used lifts]. (Research Report). Montreal, QC, Canada: Société Logique.

Reid, D. T., & Jutai, J. (1997). A pilot study of perceived clinical usefulness of a new computer-based tool for assessment of visual perception in occupational therapy practice. *Occupational Therapy International*, 4, 81-98.

Rehabilitation Engineering and Assistive Technology Society of North America [RESNA]. (1998). *RESNA resource guide for assistive technology outcomes: Assessment instruments, tools, & checklists from the field* (Vol. 2). Arlington, VA: RESNA.

Roush, S. E. (1995). The satisfaction of patients with multiple sclerosis regarding services received from physical and occupational therapy. *International Journal of Rehabilitation Health*, 1, 155-166.

Rubin, H. R. (1990). Patient evaluation of hospital care. *Medical Care*, 28, S3-S10.

Ruggeri, M. (1994). Patients' and relatives' satisfaction with psychiatric services: The state of the art of its measurement. *Social Psychiatry and Psychiatric Epidemiology*, 29, 212-227.

Ryan, S., Rigby, P., & From, W. (1996). Understanding the product needs of consumers of assistive devices. *Canadian Journal of Rehabilitation*, 9, 129-135.

Scherer, M. J. (1996a). Living in the state of stuck – How technology impacts the lives of people with disabilities (2d ed.). Cambridge, MA: Brookline.

Scherer, M. J. (1996b). Outcomes of assistive technology use on quality of life. *Disability and Rehabilitation*, 18, 439-448.

Scherer, M. J., & Vitaliti, L. T. (1997). A functional approach to technological factors and their assessment in rehabilitation. Dans S. S. Dittmar & G. E. Gresham (Eds.), *Functional assessment and outcome measures for the rehabilitation health professional* (pp. 69-88). Gaithersburg, MD: Aspen.

- Schoech, D., Cavalier, A., & Hoover, B. (1993). A model for integrating technology into a multi-agency community service delivery system. *Assistive Technology, 5*, 11-23.
- Sheikh, K. (1986). Disability scales: Assessment of reliability. *Archives of Physical Medicine and Rehabilitation, 67*, 245-249.
- Shiloh, S., Avdor, O., & Goodman, R. M. (1990). Satisfaction with genetic counseling: Dimensions and measurement. *American Journal of Medical Genetics, 37*, 522-529.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.
- Shrout, P. E., & Yager, T. J. (1989). Reliability and validity of screening scales: Effect of reducing scale length. *Journal of Clinical Epidemiology, 42*, 69-78.
- Silverstein, B., Fisher, W. P., Kilgore, K. M., Harley, J. P., & Harvey, R. F. (1992). Applying psychometric criteria to functional assessment in medical rehabilitation: II. Defining interval measures. *Archives of Physical Medicine and Rehabilitation, 73*, 507-518.
- Silverstein, B., Kilgore, K. M., Fisher, W. P., Harley, J. P., & Harvey, R. F. (1991). Applying psychometric criteria to functional assessment in medical rehabilitation: I. Exploring unidimensionality. *Archives of Physical Medicine and Rehabilitation, 72*, 631-637.
- Simon, S. E., & Patrick, A. (1997). Understanding and assessing consumer satisfaction in rehabilitation. *Journal of Rehabilitation Outcomes Measurement, 1*(5), 1-14.
- Smedley, T. C. (1990). Self-assessed satisfaction levels in elderly hearing aid, eyeglass, and denture wearers. *Ear and Hearing, 11*, 41S-47S.
- Smith, R. O. (1995). OT FACT. [Logiciel]. Rockville, MD: American Occupational Therapy Association.
- Smith, R. O. (1996a). Measuring and documenting assistive technology outcomes [Introduction]. *Assistive Technology, 8*, 69-70.

- Smith, R. O. (1996b). Measuring the outcomes of assistive technology: Challenge and innovation. *Assistive Technology*, 8, 71-81.
- Steiber, S. R., & Krowinski, W. J. (1990). *Measuring and managing patient satisfaction*. Chicago, IL: American Hospital Association.
- Strasser, S., & Davies, R. M. (1991). *Measuring patient satisfaction for improved patient services*. Ann Harbor, MI: Health Administration Press.
- Streiner, D. L., & Norman, G. R. (1989). *Health measurement scales - A practical guide to their development and use*. Oxford, NY: Oxford University Press.
- Swan, J. E., Sawyer, J. C., Van Matre, J. G., & McGee, G. W. (1985). Deepening the understanding of hospital patient satisfaction: Fulfillment and equity effects. *Journal of Health Care Marketing*, 5(3), 7-18.
- Tanner, B. A. (1982). A multi-dimensional client satisfaction instrument. *Evaluation and Program Planning*, 5, 161-167.
- Trachtman, L. (1994). Outcome measures. Are we ready to answer the tough questions? [Editorial]. *Assistive Technology*, 6, 91-92.
- Traub, R. E. (1994). *Reliability for the social sciences - Theory and applications* (vol. 3). Thousand Oaks, CA: Sage.
- Thorn, D. W., & Deitz, J. C. (1989). Examining content validity through the use of content experts. *Occupational Therapy Journal of Research*, 9, 334-346.
- Uniform Data System for Medical Rehabilitation (UDS). (1995). *Functional independence measures* (version 4.0). Buffalo, NY: Buffalo General Hospital, State University of New York.
- Valleyrand, R. J. (1989). Vers une méthodologie de validation transculturelle des questionnaires psychologiques: Implications por la recherche en langue française. *Psychologie Canadienne*, 30, 662-680.
- Vanderheiden, G. C. (1987). Service delivery mechanisms in rehabilitation technology. *American Journal of Occupational Therapy*, 41, 703-710.

- Verbrugge, L. M., & Jette, A. M. (1994). The disablement process. *Social Science Medicine*, 38, 1-14.
- Vuori, H. (1991). Patient satisfaction - Does it matter? *Quality Assurance in Health Care*, 3, 183-189.
- Ward, C. (1990). *Design for all - Consumer needs assessment project year 2 - Results of the second year of a five year study* [Research Report]. Washington, DC: Rehabilitation Engineering Center.
- Ware, J. E., Davies-Avery, A., & Stewart, A. L. (1978). The measurement and meaning of patient satisfaction. *Health and Medical Care Service Review*, 1, 1-15.
- Warren, C. G. (1993). Cost effectiveness and efficiency in assistive technology service delivery. *Assistive Technology*, 5, 61-65.
- Weiss-Lambrou, R., Demers, L., Tremblay, C., Ska, B., Lacoste, & Dansereau, J. (1997). In QUEST of user satisfaction with assistive technology devices. Dans G. Anogianakis, G. Bühler, & M. Soede (Eds.), *Advancement of Assistive Technology* (pp. 428-431). Washington, DC: IOS Press.
- Wessels, R. D., De Witte, L. P., Weiss-Lambrou, R., Demers, L., & Wijnhuizen, G. (1998). A Dutch version of QUEST (D-QUEST) applied as a routine follow-up within the service delivery process. Dans E. Placencia & E. Ballabio (Eds.), *Improving the Quality of Life for the European Citizen* (pp. 420-424). Washington, DC: IOS Press.
- Whiteneck, G. G. (1994). Measuring what matters: Key rehabilitation outcomes. *Archives of Physical Medicine and Rehabilitation*, 75, 1073-1076.
- Williams, B. (1994). Patient satisfaction: A valid concept? *Social Sciences Medicine*, 38, 509-516
- Wilson, K. G., Crupi, C. D., Greene, G., Gaulin-Jones, B., Dehoux, E., & Korol, K. T. (1995). Consumer satisfaction with a rehabilitation mobile outreach program. *Archives of Physical Medicine and Rehabilitation*, 76, 899-904.
- Winter, P. L., & Keith, R. A. (1988). A model of outpatient satisfaction in rehabilitation. *Rehabilitation Psychology*, 33, 131-142.

Woodside, A. G., Frey, L. L., & Daly, R. T. (1989). Linking service quality, customer satisfaction, and behavioral intention. *Journal of Health Care Marketing*, 9(4), 5-17.

Zastowny, T. R., Roghmann, K. L., & Cafferata, G. L. (1989). Patient satisfaction and the use of health services - Explorations in causality. *Medical Care*, 27, 705-723.

Appendice A
Formulaire d'évaluation de l'ÉSAT

FORMULAIRE D'ÉVALUATION
ÉVALUATION DE LA SATISFACTION ENVERS UNE AIDE TECHNIQUE (ÉSAT)
 Demers, L., Weiss-Lambrou, R., & Ska, B.

Nom de l'utilisateur: _____

No dossier: _____ Téléphone: _____

Adresse: _____

Nom de l'évaluateur: _____

Date de l'évaluation: _____

Aide technique évaluée (AT): _____

5. Participation à la sélection de l'aide technique:	<input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> ne s'applique pas
6. Situations de handicaps:	<input type="checkbox"/> préparation des aliments <input type="checkbox"/> prise des repas <input type="checkbox"/> hygiène corporelle <input type="checkbox"/> habillement <input type="checkbox"/> communication <input type="checkbox"/> entretien ménager <input type="checkbox"/> déplacements restreints <input type="checkbox"/> transports <input type="checkbox"/> autres:

ÉSAT partie 1: Renseignements généraux

Usager	
1. Date de naissance	(année/mois/jour)
2. Âge	_____ ans
3. Sexe:	<input type="checkbox"/> F <input type="checkbox"/> M
4. Incapacités:	<input type="checkbox"/> activités motrices <input type="checkbox"/> sens et perception <input type="checkbox"/> activités intellectuelles <input type="checkbox"/> langage <input type="checkbox"/> comportement <input type="checkbox"/> autres:

Aide technique	
7. Fréquence d'utilisation:	<input type="checkbox"/> toujours <input type="checkbox"/> fréquemment (tous les jours) <input type="checkbox"/> occasionnellement (1X/mois <-> 1X/sem) <input type="checkbox"/> rarement (<1X/mois) <input type="checkbox"/> jamais
8. Expérience avec une ou plusieurs autres ATs du même type:	<input type="checkbox"/> oui, spécifier: <input type="checkbox"/> non

<p>9. Habitude(s) de vie pour laquelle ou lesquelles l'aide technique est utilisée:</p>	<p><input type="checkbox"/> préparation des aliments</p> <p><input type="checkbox"/> prise des repas</p> <p><input type="checkbox"/> hygiène corporelle</p> <p><input type="checkbox"/> habillement</p> <p><input type="checkbox"/> communication</p> <p><input type="checkbox"/> entretien ménager</p> <p><input type="checkbox"/> déplacements restreints</p> <p><input type="checkbox"/> transports</p> <p><input type="checkbox"/> autres:</p>
<p>10. Temps écoulé depuis l'attribution de l'aide technique:</p>	<p>_____ mois</p>
<p>11. Autres aides techniques utilisées:</p>	<p>Nommer:</p>

<p>Environnement</p>	
<p>12. Type d'habitation</p>	<p><input type="checkbox"/> propriété privée</p> <p><input type="checkbox"/> logement loué</p> <p><input type="checkbox"/> résidence de personnes âgées</p> <p><input type="checkbox"/> autre:</p>
<p>13. Situation de cohabitation:</p>	<p><input type="checkbox"/> seul(e)</p> <p><input type="checkbox"/> avec conjoint(e)</p> <p><input type="checkbox"/> avec autre(s) membre(s) de la famille</p> <p><input type="checkbox"/> autre(s):</p>

<p>14. Services de maintien à domicile:</p>	<p><input type="checkbox"/> aucun</p> <p><input type="checkbox"/> organisme public:</p> <p><input type="checkbox"/> organisme privé:</p> <p><input type="checkbox"/> réseau d'entraide:</p> <p><input type="checkbox"/> ne s'applique pas</p>
<p>15. Organisme(s) responsable(s) de l'obtention de l'aide technique:</p>	<p><input type="checkbox"/> C.L.S.C.</p> <p><input type="checkbox"/> services hospitaliers de courte durée</p> <p><input type="checkbox"/> milieu de réadaptation</p> <p><input type="checkbox"/> milieu d'hébergement</p> <p><input type="checkbox"/> entreprise commerciale</p> <p><input type="checkbox"/> autre:</p>
<p>16. Entraînement reçu:</p>	<p>Description:</p>
<p>17. Délai d'attente de l'aide technique:</p>	<p>_____ semaines</p> <p>_____ mois</p>
<p>18. Coûts de l'aide technique assumés par:</p>	<p><input type="checkbox"/> RAMQ</p> <p><input type="checkbox"/> SAAQ</p> <p><input type="checkbox"/> CSST</p> <p><input type="checkbox"/> gouvernement fédéral</p> <p><input type="checkbox"/> usager (incluant la location)</p> <p><input type="checkbox"/> entourage</p> <p><input type="checkbox"/> autres :</p>

ÉSAT partie 2: Importance attribuée aux critères de satisfaction

ne sais pas
ne s'applique pas
6

pas important du tout	peu important	plus ou moins important	assez important	très important
1	2	3	4	5

ÉSAT partie 3: Cotation des critères de satisfaction

ne sais pas
ne s'applique pas
6

pas satisfait du tout	peu satisfait	plus ou moins satisfait	assez satisfait	très satisfait
1	2	3	4	5

Degré d'importance	Critères de satisfaction	Degré de satisfaction	Sources d'insatisfaction
1 2 3 4 5 6	1. La facilité d'utilisation La simplicité à utiliser l'AT.	1 2 3 4 5 6	
1 2 3 4 5 6	2. Le service de réparation et d'entretien La facilité à faire entretenir et réparer l'AT.	1 2 3 4 5 6	
1 2 3 4 5 6	3. L'entretien La facilité à maintenir soi-même l'AT propre et en bon état.	1 2 3 4 5 6	
1 2 3 4 5 6	4. L'installation La facilité à assembler et à mettre en place l'AT.	1 2 3 4 5 6	
1 2 3 4 5 6	5. L'efficacité L'atteinte du but recherché avec l'AT.	1 2 3 4 5 6	
1 2 3 4 5 6	6. Le coût Les frais occasionnés par l'achat, l'entretien et les réparations de l'AT.	1 2 3 4 5 6	
1 2 3 4 5 6	7. Le service professionnel La qualité des informations fournies sur l'AT; l'accessibilité et la compétence des professionnels.	1 2 3 4 5 6	
1 2 3 4 5 6	8. La durabilité La robustesse et la solidité de l'AT.	1 2 3 4 5 6	
1 2 3 4 5 6	9. La polyvalence La possibilité d'utiliser l'AT pour plus d'une activité (multi-usages).	1 2 3 4 5 6	

ne sais pas
ne s'applique pas
6

pas important du tout 1
peu important 2
plus ou moins important 3
assez important 4
très important 5

ne sais pas
ne s'applique pas
6

pas satisfait du tout 1
peu satisfait 2
plus ou moins satisfait 3
assez satisfait 4
très satisfait 5

Degré d'importance		Critères de satisfaction	Degré de satisfaction	Sources d'insatisfaction	
1	2 3 4 5 6	10. Les ajustements Le degré de facilité à régler et à ajuster les composantes de l'AT.	1 2 3 4 5 6		
1	2 3 4 5 6	11. Le confort Le bien-être physique et psychologique associé à l'utilisation de l'AT.	1 2 3 4 5 6		
1	2 3 4 5 6	12. L'attribution Le degré de facilité à acquérir l'AT incluant le temps requis.	1 2 3 4 5 6		
1	2 3 4 5 6	13. Le suivi des services La qualité des services post-attribution de l'AT.	1 2 3 4 5 6		
1	2 3 4 5 6	14. L'apparence L'aspect de l'AT, sa conception, sa forme, sa couleur et son acceptabilité.	1 2 3 4 5 6		
1	2 3 4 5 6	15. Le transport La facilité de déplacement par le mode de transport souhaité.	1 2 3 4 5 6		
1	2 3 4 5 6	16. La compatibilité technique L'intégration de l'AT avec les autres technologies utilisées.	1 2 3 4 5 6		
1	2 3 4 5 6	17. Le poids La facilité à soulever et/ou à déplacer l'AT.	1 2 3 4 5 6		
1	2 3 4 5 6	18. La sécurité La sûreté, la fiabilité et l'absence de risque de l'AT.	1 2 3 4 5 6		
1	2 3 4 5 6	19. Les dimensions La commodité du format de l'AT (grosueur, largeur, longueur).	1 2 3 4 5 6		

ne sais pas
ne s'applique pas
6

pas important
du tout 1

peu important 2

plus ou moins important 3

assez important 4

très important 5

ne sais pas
ne s'applique pas
6

pas satisfait
du tout 1

peu satisfait 2

plus ou moins satisfait 3

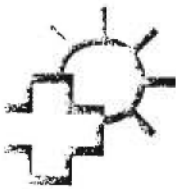
assez satisfait 4

très satisfait 5

Degré d'importance		Critères de satisfaction	Degré de satisfaction	Sources d'insatisfaction
1	2 3 4 5 6	20. La motivation L'intérêt pour l'AT pour l'utiliser chez soi et dans les endroits publics.		
1	2 3 4 5 6	21. L'entraînement L'apprentissage requis pour utiliser l'AT.		
1	2 3 4 5 6	22. Le soutien de l'entourage Le support physique ou psychologique de la famille, des pairs et de l'employeur pour utiliser l'AT.		
1	2 3 4 5 6	23. La réaction des autres L'attitude positive et encourageante de l'entourage.		
1	2 3 4 5 6	24. L'effort Le peu de fatigue physique ou psychologique associé à l'AT.		
1	2 3 4 5 6	25. Autre:	1 2 3 4 5 6	
1	2 3 4 5 6	26. Autre:	1 2 3 4 5 6	
1	2 3 4 5 6	27. Autre:	1 2 3 4 5 6	
		SATISFACTION GLOBALE	1 2 3 4 5 6	

Ne peut être reproduit ni en partie ni en totalité sans la permission des auteurs.
04/1997/QC (version expérimentale 7)

Appendice B
Certificat d'éthique



CENTRE
HOSPITALIER
CÔTE-DES-NEIGES

CENTRE
D'ACCUEIL
ALFRED-DESROCHERS

**RAPPORT DE L'ÉTABLISSEMENT
OÙ LA RECHERCHE SERA ENTREPRISE**

Centre de
gériatrie affilié à
Université de Montréal

Le comité d'éthique de la recherche du Centre hospitalier Côte-des-Neiges, a examiné la demande pour le projet intitulé:

L'évaluation de la satisfaction envers une aide technique (ESAT):
étude de fidélité et de validité (réf: 96.11.01)

Présenté par: Bernadette Ska

Date: 25 novembre 1996

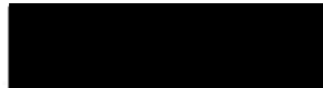
et juge la recherche faisant appel à des sujets humains acceptable au point de vue éthique.

=====

LE COMITÉ D'ÉTHIQUE DE LA RECHERCHE EST COMPOSÉ DES MEMBRES SUIVANTS:



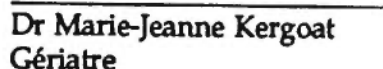
Dr Céline Crowe, D.S.P.
Présidente du comité



Louise Francoeur
Infirmière clinicienne



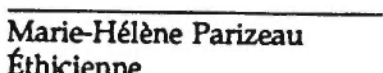
Dr Nathalie Champoux
Omnipraticienne



Dr Marie-Jeanne Kergoat
Gériatre



Guylaine Ferland
Chercheure en nutrition



Marie-Hélène Parizeau
Éthicienne



Francine Fontaine
Neuropsychologue



biendaire
prix
-Lachapelle



excellence
es sociales
989

4565, chemin Queen-Mary
Montréal (Québec) H3W 1W5
(514) 340-1424
Fax: (514) 340-3500

5325, avenue Victoria
Montréal (Québec) H3W 2P2
(514) 731-3891
Fax: (514) 340-3500