

JEAN-SÉBASTIEN RENAUD

**LA MESURE DE LA SATISFACTION DES CLIENTS :
Effet de la polarité, du nombre de points et de l'étiquetage de
l'échelle de réponse sur la distribution des réponses et sur les
qualités psychométriques**

Thèse présentée
à la Faculté des études supérieures et postdoctorales de l'Université Laval
dans le cadre du programme de doctorat en administration et évaluation en éducation
pour l'obtention du grade de *Philosophiae Doctor* (Ph. D.)

DÉPARTEMENT DES FONDEMENTS ET PRATIQUES EN ÉDUCATION
FACULTÉ DES SCIENCES DE L'ÉDUCATION
UNIVERSITÉ LAVAL
QUÉBEC

2012

Résumé

Plusieurs auteurs suspectent que la grande proportion de scores élevés dans les études de satisfaction des clients restreint la variance et les covariances des items et d'un questionnaire de satisfaction et, par conséquent, que ceci affecte négativement leurs qualités psychométriques. Dans ce contexte, des auteurs ont suggéré d'accroître la sensibilité des échelles de réponse des questionnaires de satisfaction afin d'augmenter la variance associée à la caractéristique mesurée (*i.e.*, la variance vraie), c'est-à-dire la satisfaction. Pour ce faire, ils ont proposé de manipuler deux caractéristiques de l'échelle de réponse : sa polarité et son nombre de points.

S'inscrivant dans cette problématique, l'objectif de cette étude est d'évaluer l'effet de la polarité (bipolaire c. unipolaire), du nombre de points (six c. dix) et, à titre exploratoire, de l'étiquetage (*i.e.*, avec ou sans étiquettes) de l'échelle de réponse sur la distribution des réponses ainsi que sur les qualités psychométriques d'un questionnaire de satisfaction des clients et de ses items. L'étude prend comme cadre théorique l'approche cognitive des méthodes de sondage, le principal courant d'études s'intéressant au comportement de réponse à un questionnaire, et emploie un devis quasi-expérimental, plus précisément un plan factoriel incomplet. Ce plan comporte six groupes de comparaison et approximativement 150 sujets se retrouvent dans chacun d'eux, pour une taille totale d'échantillon de 907 participants.

Les résultats de l'étude indiquent que les indices de dispersion (*i.e.*, écart-type, variance) et les qualités psychométriques des items et d'un questionnaire de satisfaction sont égaux ou inférieurs lorsque l'échelle de réponse est unipolaire plutôt que bipolaire. Conséquemment, nous recommandons d'employer une échelle de réponse bipolaire plutôt qu'unipolaire dans un questionnaire de satisfaction. Les résultats nous apprennent aussi que l'effet de la polarité est difficile à prédire, car il diffère selon que l'échelle de réponse compte six ou dix

points et selon qu'elle possède ou pas des étiquettes. En ce qui a trait à l'effet du nombre de points de l'échelle de réponse, les résultats de l'étude amènent à conclure qu'utiliser une échelle de réponse en six points plutôt qu'en dix est sans effet sur les qualités psychométriques d'un questionnaire de satisfaction et de ses items. L'échelle de réponse en six points pourrait donc représenter un choix parcimonieux. L'étude fait aussi ressortir que la polarité, le nombre de points et l'étiquetage de l'échelle de réponse ont peu d'effet sur les qualités psychométriques d'un questionnaire de satisfaction et de ses items et, en ce sens, que ces caractéristiques de l'échelle de réponse ont une importance secondaire dans la construction d'un questionnaire de satisfaction. Enfin, les résultats permettent de penser que la distribution asymétrique négative des scores de satisfaction ne représente pas une menace importante pour les qualités psychométriques d'un questionnaire de satisfaction et de ses items.

Mots clés : questionnaire autoadministré, satisfaction des clients, psychométrie, échelle de réponse, items, qualités psychométriques

Je désire prendre ces quelques lignes pour remercier les personnes qui ont joué un rôle important pour moi au cours de ces années d'étude au doctorat.

Je tiens à commencer en remerciant mon directeur de thèse, Pierre Valois, une personne que j'apprécie et envers qui j'ai la plus haute estime. Cela fait maintenant huit ans que j'ai la chance de te connaître et d'apprendre de toi, dont les grandes qualités de chercheur ne sont plus à prouver. Pendant ces années tu as su être un guide bienveillant et aidant. Tu m'as ouvert plusieurs portes et as contribué de manière importante au développement de mes compétences. Je veux donc t'exprimer en ces quelques lignes, Pierre, toute ma gratitude.

Je me dois aussi de remercier le Musée de la civilisation de Québec, et tout particulièrement Lucie Daignault, du Service de la recherche et de l'évaluation du musée, pour m'avoir ouvert la porte et m'avoir donné de leurs conseils et de leur temps. Dans le même ordre d'idées, je remercie Richard Bédard et Geneviève Provencher-St-Cyr, qui m'ont aidé avec la collecte de données, de même que les participants de l'étude. Votre précieuse collaboration a été un facteur de succès dans la réalisation de cette étude et je vous en suis reconnaissant.

Je voudrais de plus remercier Jean Robitaille, professeur en sciences de la consommation, sur qui je peux compter et qui a cru en moi depuis le baccalauréat. Tu as toujours été prêt à m'accorder soutien et temps, ce qui a été très précieux pour moi. Également du département d'économie agroalimentaire et des sciences de la consommation, je remercie les professeures Marie Lachance et Gale West, qui ont été les premières personnes à me donner la chance de travailler en recherche scientifique et qui ont éveillé chez moi cet intérêt pour la recherche.

J'en profite aussi pour mentionner à quel point il a été agréable de pouvoir côtoyer et tisser des liens avec les personnes « du 4^{ième} » de la tour de la Faculté des sciences de l'éducation. Ça a été et c'est toujours un réel plaisir de vous voir, autant ceux qui n'y sont plus aujourd'hui que ceux qui y sont toujours. Parmi ces personnes, j'aimerais remercier

Éric Frenette, que j'ai connu lorsqu'il était étudiant au doctorat et qui est maintenant professeur à cette même faculté. Ta disponibilité et tes coups de pouce occasionnels ont été très appréciés.

Je ne peux surtout pas passer sous silence le soutien que j'ai reçu de ma famille. Je tiens à remercier ma conjointe, Marie-Ève, dont le soutien au quotidien a été fort apprécié. Tu as su m'écouter, m'aider et m'encourager, et ce, malgré le fait que la poursuite d'études doctorales empiète sur la vie et les projets de couple. Merci d'avoir cru en moi et de m'avoir appuyé dans ma décision d'aller jusqu'au bout de ce rêve. Je remercie mes parents, Jean-Claude et Françoise, qui m'ont appuyé et encouragé du mieux qu'ils le pouvaient dans toute cette aventure. Je remercie aussi mon parrain, Michel, qui a toujours tenu à contribuer à faire en sorte que ce projet soit un succès.

Je termine avec mes amis proches, des personnes très importantes pour moi. Je pense ici à Jean-Denis, Simon, Vincent, Jean-François, Daniel et Michael. Je vous remercie de vos encouragements. Votre précieuse amitié a su perdurer au cours de toutes ces années, même si parfois on ne se voyait pas aussi souvent que j'aurais aimé étant donné que je conjuguais études et travail.

Table des matières

Introduction.....	1
Chapitre 1 : Problématique	3
La satisfaction des clients : état des lieux	3
La satisfaction des clients au centre des préoccupations de l'organisation	3
La définition du concept de satisfaction	5
L'opérationnalisation et la mesure du concept de satisfaction des clients	8
Le problème de recherche.....	11
Objectif général et pertinence de l'étude	22
Chapitre 2 : Cadre théorique.....	24
Le modèle théorique de la satisfaction des clients.....	24
L'approche cognitive des méthodes de sondage.....	30
La compréhension de la question.....	32
La récupération de l'information pertinente en mémoire	33
Le jugement	33
La réponse.....	34
Les études sur la relation entre la polarité de l'échelle de réponse et la distribution des réponses.....	37
Les études sur la relation entre le nombre de points de l'échelle de réponse et les qualités psychométriques d'un outil de mesure	50
Objectif spécifique et hypothèses de recherche.....	56

Chapitre 3 : Méthodologie	61
Variables et devis de recherche	61
Variables dépendantes	62
Variables indépendantes	63
Devis de recherche.....	66
Participants.....	67
Questionnaire.....	70
Questions sur le profil des participants.....	71
Les instruments de mesure et leur élaboration.....	73
Plan d'analyse des données.....	107
Chapitre 4 : Résultats	114
Caractéristiques psychométriques des outils de mesure utilisés dans l'étude	114
Mise à l'épreuve des hypothèses de recherche	118
Hypothèses 1 à 9 : effet de la polarité de l'échelle de réponse	118
Hypothèses 10 à 15 : effet du nombre de points dans l'échelle de réponse	148
Mise à l'épreuve des 15 hypothèses de recherche : un résumé	161
Chapitre 5 : Discussion.....	164
Effet de la polarité de l'échelle de réponse (H_1 à H_9).....	164
Effet du nombre de points de l'échelle de réponse (H_{10} à H_{15})	171
Constats et implications.....	175
Limites de l'étude	180
Conclusion	184

Bibliographie	187
Annexes	203
Annexe 1 : Définitions de la satisfaction des clients.....	204
Annexe 2 : Échelle de satisfaction des clients de Oliver (1980, 1997) traduite en français par Vanhamme (2002).....	211
Annexe 3 : Directives pour réaliser la collecte de données au Musée de la civilisation.....	213
Annexe 4 : Attributs d'une exposition muséale associés à la satisfaction des visiteurs selon la littérature récente	216
Annexe 5 : Questionnaire de l'étude qualitative exploratoire.....	226
Annexe 6 : Réponses des participants au questionnaire de l'étude qualitative exploratoire	231
Annexe 7 : Questionnaire de l'étude qualitative confirmatoire	234
Annexe 8 : Première version de la mesure multi-items de satisfaction (72 items).....	240
Annexe 9 : Première version de la mesure de l'attitude à l'égard de la fréquentation du Musée de la civilisation (26 items).....	244
Annexe 10 : Questionnaire des entrevues cognitives.....	246
Annexe 11 : Questionnaire utilisé pour la mise à l'épreuve des qualités psychométriques des cinq outils de mesure développés.....	255
Annexe 12 : Questionnaire de l'étude	263
Annexe 13 : Évaluation de la validité nomologique de la mesure de la satisfaction des clients : indices d'ajustement et coefficient de détermination du modèle structurel	272

Liste des tableaux

Tableau 1.1	Lien entre la moyenne et la variance pour un item dichotomique (tiré de Courville, 2004)	16
Tableau 2.1	Résumé des composantes du processus de réponse à une question de sondage (adapté de Tourangeau <i>et al.</i> , 2000, p. 8).....	31
Tableau 3.1	Nombre de participants dans chacun des six groupes de comparaison.....	69
Tableau 3.2	Répartition par catégorie d'âge et par sexe des 907 sujets de l'échantillon.....	69
Tableau 3.3	Devis de collecte des données de la deuxième étape de l'élaboration de la mesure multi-items de satisfaction des visiteurs (étude qualitative exploratoire).....	77
Tableau 3.4	Corrélations entre les variables du modèle théorique de la satisfaction des clients lors de l'analyse d'items préliminaire ($n = 65$)	102
Tableau 3.5	Résultats des analyses de fidélité réalisées avec les données du prétest ($n = 65$)	103
Tableau 3.6	Plan factoriel $2 \times 2 \times 2$ incomplet	109
Tableau 4.1	Score global et indices psychométriques des deux outils de mesure de la satisfaction des clients pour chacune des six conditions testées.....	116
Tableau 4.2	Score global et indices psychométriques des outils de mesure de l'attitude envers la fréquentation du musée et des intentions comportementales.....	117
Tableau 4.3	Valeurs moyennes des indices de difficulté des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points).....	120
Tableau 4.4	ANOVA testant l'effet de la polarité et de l'étiquetage sur les indices de difficulté des items de satisfaction (échelles de réponse en six points)	120
Tableau 4.5	Analyses de contrastes évaluant l'effet de la polarité sur les indices de difficulté des items de satisfaction selon l'étiquetage (échelles de réponse en six points)	121

Tableau 4.6	ANOVA testant l'effet de la polarité sur les indices de difficulté des items de satisfaction (échelles de réponse en dix points).....	122
Tableau 4.7	Valeurs moyennes des écarts-types des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points).....	123
Tableau 4.8	ANOVA testant l'effet de la polarité et de l'étiquetage sur les écarts-types des items de satisfaction (échelles de réponse en six points).....	124
Tableau 4.9	Analyses de contrastes évaluant l'effet de la polarité sur les écarts-types des items de satisfaction selon l'étiquetage (échelles de réponse en six points)	124
Tableau 4.10	ANOVA testant l'effet de la polarité sur les écarts-types des items de satisfaction (échelles de réponse en dix points)	126
Tableau 4.11	Valeurs moyennes des indices d'asymétrie des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points).....	127
Tableau 4.12	ANOVA testant l'effet de la polarité et de l'étiquetage sur les indices d'asymétrie des items de satisfaction (échelles de réponse en six points)	128
Tableau 4.13	Analyses de contrastes évaluant l'effet de la polarité sur les indices d'asymétrie des items de satisfaction selon l'étiquetage (échelles de réponse en six points)	128
Tableau 4.14	ANOVA testant l'effet de la polarité sur les indices d'asymétrie des items de satisfaction (échelles de réponse en dix points).....	130
Tableau 4.15	Score global de satisfaction en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)	131
Tableau 4.16	ANOVA testant l'effet de la polarité et de l'étiquetage sur le score global de satisfaction (échelles de réponse en six points).....	131
Tableau 4.17	Analyses de contrastes évaluant l'effet de la polarité sur le score global de satisfaction selon l'étiquetage (échelles de réponse en six points)	132
Tableau 4.18	ANOVA testant l'influence de la polarité sur le score global de satisfaction (échelles de réponse en dix points)	133
Tableau 4.19	Variance du score global de satisfaction en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)	135

Tableau 4.20	Variance du score global de satisfaction en fonction de la polarité (échelles de réponse en dix points)	135
Tableau 4.21	Moyennes des indices de discrimination des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points).....	137
Tableau 4.22	ANOVA testant l'effet de la polarité et de l'étiquetage sur les indices de discrimination des items de satisfaction (échelles de réponse en six points)	137
Tableau 4.23	Analyses de contrastes évaluant l'effet de la polarité sur les indices de discrimination des items de satisfaction selon l'étiquetage (échelles de réponse en six points)	137
Tableau 4.24	ANOVA testant l'effet de la polarité sur les indices de discrimination des items de satisfaction (échelles de réponse en dix points).....	139
Tableau 4.25	Coefficients alpha de Cronbach selon la polarité et de l'étiquetage (échelles de réponse en six points)	140
Tableau 4.26	Coefficients alpha de Cronbach selon la polarité de l'échelle de réponse (échelles de réponse en dix points)	141
Tableau 4.27	Indices de validité de convergence en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)	142
Tableau 4.28	Indices de validité de convergence en fonction de la polarité (échelles de réponse en dix points).....	143
Tableau 4.29	Corrélations entre la satisfaction des clients et les variables associées en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)	144
Tableau 4.30	Corrélations entre la satisfaction des clients et les variables associées en fonction de la polarité (échelles de réponse en dix points).....	144
Tableau 4.31	Indices de validité nomologique en fonction de la polarité selon l'étiquetage (échelles de réponse en six points)	146
Tableau 4.32	Tests de l'invariance des trois paramètres structurels en fonction de la polarité selon l'étiquetage (échelles de réponse en six points).....	146
Tableau 4.33	Indices de validité nomologique en fonction de la polarité (échelles de réponse en dix points).....	147

Tableau 4.34	Tests de l'invariance des trois paramètres structurels selon la polarité (échelles de réponse en dix points)	147
Tableau 4.35	Coefficients de variation moyens des items de satisfaction selon le nombre de points et la polarité de l'échelle de réponse.....	150
Tableau 4.36	ANOVA testant l'effet du nombre de points et de la polarité de l'échelle de réponse sur les coefficients de variation des items de satisfaction.....	150
Tableau 4.37	Analyses de contrastes évaluant l'effet du nombre de points sur les coefficients de variation des items de satisfaction selon la polarité de l'échelle de réponse	151
Tableau 4.38	Variance du score global de satisfaction en fonction du nombre de points de réponse selon que l'échelle de réponse est bipolaire ou unipolaire	153
Tableau 4.39	Indices de discrimination moyens en fonction du nombre de points et de la polarité de l'échelle de réponse	154
Tableau 4.40	ANOVA testant l'effet du nombre de points et de la polarité de l'échelle de réponse sur les indices de discrimination des items de satisfaction	154
Tableau 4.41	Analyses de contrastes évaluant l'effet du nombre de points de réponse sur les indices de discrimination des items de satisfaction en fonction de la polarité de l'échelle de réponse	155
Tableau 4.42	Coefficients alpha de Cronbach en fonction du nombre de points de l'échelle de réponse selon sa polarité	156
Tableau 4.43	Indices de validité de convergence en fonction du nombre de points et de la polarité de l'échelle de réponse.....	158
Tableau 4.44	Corrélations entre la satisfaction des clients et les variables associées en fonction du nombre de points et de la polarité de l'échelle de réponse	159
Tableau 4.45	Paramètres structurels et coefficient de détermination selon le nombre de points et la polarité de l'échelle de réponse.....	160
Tableau 4.46	Tests de l'invariance des trois paramètres structurels selon le nombre de points de réponse pour les échelles de réponse bipolaires et unipolaires.....	160
Tableau 4.47	Décisions pour les hypothèses relatives à l'effet de la polarité de l'échelle de réponse (H_1 à H_9)	162

Tableau 4.48	Décisions pour les hypothèses relatives à l'effet du nombre de points dans l'échelle de réponse (H_{10} à H_{15})	163
Tableau A1.1	Définitions les plus courantes du concept de satisfaction des clients.....	205
Tableau A2.1	Échelle de satisfaction des clients de Oliver (1980, 1997) traduite en français par Vanhamme (2002)	212
Tableau A4.1	Attributs d'une exposition muséale associés à la satisfaction des visiteurs.....	217
Tableau A8.1	Les 72 items de la première version de la mesure multi-items de satisfaction des visiteurs envers les expositions.....	241
Tableau A13.1	Indices d'ajustement et coefficient de détermination selon la polarité et l'étiquetage (échelles de réponse en 6 points).....	273
Tableau A13.2	Indices d'ajustement et coefficient de détermination selon la polarité (échelles de réponse en 10 points).....	273
Tableau A13.3	Indices d'ajustement et coefficient de détermination selon le nombre de points de l'échelle de réponse.....	273

Liste des figures

Figure 1.1	La distribution dissymétrique négative des scores à un questionnaire de satisfaction des clients (adaptée de Peterson & Wilson, 1992).....	12
Figure 1.2	L'évolution de l' <i>American Customer Satisfaction Index</i> pour la période 1994-2009 (tirée du site Internet de l'ACSI).....	12
Figure 1.3	Relation entre la moyenne des scores de satisfaction (exprimés en pourcentage) et l'ampleur de l'écart-type (tirée de Cummins, 2003).....	17
Figure 1.4	Illustration de la première proposition : utiliser une échelle de réponse unipolaire plutôt que bipolaire.....	20
Figure 1.5	Illustration de la deuxième proposition : augmenter le nombre de points dans l'échelle de réponse.....	21
Figure 2.1	Représentation circulaire (dynamique) du modèle théorique de la satisfaction des clients d'Oliver (1980, 1997).....	25
Figure 2.2	Le modèle théorique de la satisfaction des clients sur lequel repose l'étude.....	30
Figure 2.3	Les cinq échelles de réponse de l'étude de Klockars et Yamagishi (1988).....	41
Figure 3.1	L'échelle de réponse bipolaire en six points (avec des étiquettes).....	64
Figure 3.2	L'échelle de réponse unipolaire en six points (avec des étiquettes).....	64
Figure 3.3	L'échelle de réponse bipolaire en six points (sans étiquettes).....	65
Figure 3.4	L'échelle de réponse unipolaire en six points (sans étiquettes).....	65
Figure 3.5	L'échelle de réponse bipolaire en dix points.....	65
Figure 3.6	L'échelle de réponse unipolaire en dix points.....	65
Figure 3.7	Validité nomologique : modèle structurel des relations théoriques entre la satisfaction des clients, l'attitude envers le comportement et les intentions comportementales.....	111

Figure 4.1	Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les indices de difficulté des items de satisfaction (échelles de réponse en six points)	121
Figure 4.2	Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les écarts-types des items de satisfaction (échelles de réponse en six points)	125
Figure 4.3	Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les indices d'asymétrie des items de satisfaction (échelles de réponse en six points)	129
Figure 4.4	Illustration de l'effet d'interaction entre la polarité et l'étiquetage des points intermédiaires de l'échelle de réponse sur le score global de satisfaction (échelles de réponse en six points)	132
Figure 4.5	Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les indices de discrimination des items de satisfaction (échelles de réponse en six points)	138
Figure 4.6	Illustration de l'effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse sur les coefficients de variation des items de satisfaction	151
Figure 4.7	Illustration de l'effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse sur les indices de discrimination des items de satisfaction	155

Introduction

La satisfaction des clients est un concept central en sciences de la gestion, notamment en marketing relationnel. La recherche sur ce concept est relativement récente, les premiers travaux datant des années 60. Depuis, le concept de satisfaction des clients a été adopté dans plusieurs autres champs de recherche, tels que l'éducation (p. ex. la satisfaction des étudiants), les sciences de la santé (p. ex. la satisfaction des patients) et la psychologie (p. ex. la satisfaction envers sa vie, son couple, ses relations sociales, etc.), pour ne nommer que ceux-ci. Il s'agit maintenant d'un concept incontournable dès que l'on étudie la relation entre une organisation ou une personne qui fournit des produits ou des services et les individus qui les consomment (*e.g.*, entreprises/clients, organisations gouvernementales/citoyens, hôpitaux/patients, institutions scolaires/étudiants). Indépendamment du type d'organisation, l'intérêt d'évaluer la satisfaction des clients est double. Dans un premier temps, il tient au fait qu'elle est associée à l'adoption chez les clients de comportements qui sont recherchés par l'organisation. Ces comportements recherchés diffèrent grandement selon le contexte, allant du comportement de rachat (*i.e.*, la fidélisation des clients) à une meilleure adhérence du patient à son traitement (Ruggeri, 1994). Dans un deuxième temps, il consiste en ce que le degré de satisfaction des clients est utilisé comme indicateur de la qualité des produits ou des services offerts. Ce faisant, l'évaluation de la satisfaction des clients sert aux fournisseurs de produits et de services à établir leurs priorités en vue d'améliorer ce qu'ils offrent.

Une des principales préoccupations soulevées dans la littérature sur la mesure de la satisfaction des clients est que les scores de satisfaction sont généralement élevés, ce qui s'observe par l'asymétrie négative de leur distribution. Plusieurs auteurs suspectent que ce phénomène restreint la variance et les covariances des items et d'un questionnaire de satisfaction et, par conséquent, ceci affecte négativement leurs qualités psychométriques. Dans ce contexte, des auteurs ont suggéré d'accroître la sensibilité des échelles de réponse des questionnaires de satisfaction afin d'augmenter la variance associée à la caractéristique

mesurée (*i.e.*, la variance vraie), c'est-à-dire la satisfaction. Pour ce faire, ils ont proposé de manipuler deux caractéristiques de l'échelle de réponse : sa polarité et son nombre de points.

S'inscrivant dans cette problématique, l'objectif de cette étude est d'évaluer l'effet de la polarité (bipolaire c. unipolaire), du nombre de points (six c. dix) et, à titre exploratoire, de l'étiquetage (*i.e.*, avec ou sans étiquettes) de l'échelle de réponse sur la distribution des réponses ainsi que sur les qualités psychométriques d'un questionnaire de satisfaction des clients et de ses items. Afin d'atteindre cet objectif, l'étude prend comme cadre théorique l'approche cognitive des méthodes de sondage, le principal courant d'études s'intéressant au comportement de réponse à un questionnaire. Le devis de recherche, de type quasi-expérimental, est un plan factoriel incomplet à six conditions. L'échantillon, quant à lui, est composé de 907 participants.

La présente thèse rend compte de cette étude en cinq chapitres. Le premier chapitre cerne la problématique de l'étude et se termine avec la présentation de son objectif général. Le second chapitre présente le cadre théorique sur lequel s'appuie la recherche et prend fin avec l'objectif spécifique et les hypothèses de recherche. Le troisième chapitre décrit la méthodologie employée. Le quatrième chapitre présente les résultats de l'étude. Le cinquième chapitre discute des résultats, ce après quoi nous concluons cette thèse.

Chapitre 1 : Problématique

Ce premier chapitre cerne la problématique de cette étude. Il débute en dressant un état des lieux sur le concept de satisfaction des clients pour enchaîner ensuite avec la présentation du problème de recherche. Il prend fin avec la présentation de l'objectif général et de la pertinence de l'étude.

La satisfaction des clients : état des lieux

L'état des lieux au sujet du concept de satisfaction des clients consiste, d'une part, à mettre en perspective les connaissances issues de la littérature et, d'autre part, à mieux situer le lecteur quant à notre sujet de recherche.

La satisfaction des clients au centre des préoccupations de l'organisation

Le concept de satisfaction des clients n'a pas toujours eu l'importance qu'on lui accorde aujourd'hui. En effet, la gestion de la mise en marché des produits et des services a beaucoup évolué au 20^e siècle, et ce, tant dans les secteurs privé que public. Par exemple, avant les années 60, la gestion de cette mise en marché était centrée sur le produit ou le service, c'est-à-dire l'extrant de l'organisation. En conséquence, la qualité ou la valeur d'un produit ou d'un service était évaluée par l'organisation sur la base de ce qu'elle croyait être un bon produit ou service. C'est au cours des années 60, avec l'arrivée d'un nouveau paradigme de gestion, que le concept de satisfaction des clients devient important aux yeux des organisations et qu'il transforme la façon d'entrevoir la mise en marché des biens et services. Désormais, le principal évaluateur de la qualité ou de la valeur d'un produit ou

d'un service est le client lui-même (Keegan & Seringhaus, 1999). De cette façon, l'organisation s'assure que ce qu'elle offre satisfait ses clients. De même, elle augmente également ses chances de survie, ce qui est capital dans une conjoncture où la concurrence est féroce. En effet, pour les organisations, un degré élevé de satisfaction des clients est associé à une augmentation du taux de rétention des clients, à une augmentation de la rentabilité, à une meilleure compétitivité et, enfin, à une augmentation des parts de marché (Vavra, 1997)¹. À titre d'exemple, Anderson, Fornell et Lehman (1994) ont trouvé dans leur étude un lien entre la satisfaction des clients et la performance financière de l'entreprise. Ces chercheurs ont estimé, à l'aide des données du *Swedish Customer Satisfaction Barometer*, qu'une firme moyenne en Suède, c'est-à-dire qui bénéficie d'un retour sur investissement de 10,83 % par an, devrait faire passer ce gain financier annuel à 11,54 % dans cinq ans si elle augmente son score de satisfaction des clients d'un pourcent par année pendant cette période.

Ainsi, au fil des années, les principes modernes de gestion de la mise en marché des produits et services se sont construits sur le principe qu'il faut satisfaire les clients (Cooper, Cooper, & Duban, 1989). Conséquemment, la satisfaction des clients a acquis un statut de concept fondamental et central en gestion de la mise en marché des produits et des services (Yi, 1990). Déjà au début des années 90, plus de 15 000 articles scientifiques et non scientifiques avaient été publiés concernant la satisfaction des clients (Peterson & Wilson, 1992). Satisfaire les clients et même augmenter leur degré de satisfaction est devenu un objectif prioritaire pour les organisations (Kotler, Armstrong, & Cunningham, 1999; Wirtz, 2001). De plus en plus, la tendance est de mettre la satisfaction des clients au centre des priorités de l'entreprise (Koskas, 1997). Autrement dit, ce n'est plus seulement la gestion de la mise en marché qui est orientée vers l'objectif de satisfaire les clients, mais toute la gestion de l'organisation qui intègre cet objectif. À titre d'indicateur de la qualité de l'extrait produit par l'organisation (Alain, 2002), la satisfaction des clients devient du même coup un indicateur de la performance de l'organisation. L'évaluation de la

¹ Malheureusement, Vavra (1997) appuie ses propos sur très peu de publications scientifiques, c'est-à-dire évaluées par les pairs.

satisfaction des clients permet alors aux gestionnaires d'établir les priorités en matière d'améliorations pour leur organisation.

À mesure que cette nouvelle logique de gestion s'est répandue, les évaluations de la satisfaction des clients se sont multipliées et sont maintenant devenues chose courante (Koskas, 1997). Cette façon de faire a même été adoptée par les organisations qui ne sont pas dans un contexte de compétition. Par exemple, les gouvernements du Canada et du Québec évaluent la satisfaction de leurs clients afin d'évaluer la qualité de leurs services et la performance de leur organisation (*e.g.*, CEFRIO, 2003; Neill, Tremblay, Végiard, Lavoie, & Moisan, 2007; Phoenix Strategic Perspectives Inc., 2006; Schmidt & Strickland, 1998). La mesure de la satisfaction des clients est d'ailleurs une composante intégrale de bon nombre de travaux d'évaluation de programme (Worthington, 2005). Ainsi, depuis le début des années 2000 au gouvernement du Québec, « l'évaluation de la satisfaction est au cœur même de la mission et des responsabilités imputées aux diverses instances impliquées dans l'organisation des services de santé et des services sociaux » (Neill *et al.*, 2007, p. 11). En somme, la satisfaction des clients est devenue aujourd'hui un concept capital qui occupe beaucoup de place, et ce, tant dans la littérature scientifique que dans la gestion des organisations.

La définition du concept de satisfaction

Selon Oliver (1997), le terme satisfaction serait dérivé du latin « *satis* » (*i.e.*, assez) et « *facere* » (*i.e.*, faire). Toujours selon cet auteur, cette origine étymologique signifie qu'un produit ou un service satisfaisant est capable d'apporter au client ce qu'il cherchait dans une quantité qu'il juge suffisante. Parker et Mathews (2001) ajoutent que le terme satisfaction possède des similarités avec d'autres mots tels que « satisfaisant », « satisfaire » et « satiacion ».

Toutefois, les définitions scientifiques de la satisfaction des clients s'éloignent de son origine latine. Les premières références à la satisfaction des clients datent de la deuxième moitié du vingtième siècle. En effet, l'étude de la satisfaction des clients est née avec l'arrivée du marketing et est donc plutôt récente (années 60). Cardozo (1965) est sans doute un des tout premiers auteurs à avoir tenté de comprendre en quoi consiste la satisfaction des clients. Sans la définir, il conclut son article en indiquant qu'il s'agit probablement d'un concept global, plus grand que la simple évaluation d'un produit. Il suggère qu'il pourrait s'agir de l'évaluation de l'expérience de consommation dans son ensemble, incluant le produit. La définition du concept de satisfaction des clients généralement identifiée par les auteurs comme étant la première à être apparue dans la littérature est celle de Howard et Sheth (1969) dans leur livre « *The Theory of Buyer Behavior* » (p. 145) : « l'impression (état cognitif) qu'a l'acheteur d'avoir été adéquatement ou non récompensé pour les sacrifices qu'il a faits », traduction libre de « *the buyer's cognitive state of being adequately or inadequately rewarded for the sacrifices he has undergone* ». Cette définition ainsi que plusieurs autres sont présentées dans le tableau A1.1 de l'annexe 1. Ce dernier présente par ordre chronologique les définitions du concept de satisfaction des clients les plus populaires. Cette liste de définitions est inspirée, entre autres, des recensions faites par Giese et Cote (2000) et Vanhamme (2002).

L'examen du tableau A1.1 permet de constater qu'il existe une panoplie de définitions de la satisfaction des clients. Le plus souvent, la satisfaction des clients y est conceptualisée comme un jugement ou une évaluation à l'égard d'un produit ou d'un service reçu (Anderson, Fornell, & Lehmann, 1994; Bloemer & Kasper, 1995; Day, 1984; Engel, Kollat, & Blackwell, 1968; Fornell, 1992; Hunt, 1977; LaBarbera & Mazursky, 1983; Mano & Oliver, 1993; Oliver, 1981, 1997; Ostrom & Iacobucci, 1995; Swan, Trawick, & Carrol, 1980; Westbrook, 1987; Westbrook & Oliver, 1991). Un peu moins fréquemment, la satisfaction des clients y est aussi conceptualisée comme une réponse affective associée à l'évaluation du bien ou service reçu (Halstead, Hartman, & Schmidt, 1994; Westbrook & Reilly, 1983; Woodruff, Cadotte, & Jenkins, 1983). À notre connaissance, la définition de la satisfaction la plus populaire, c'est-à-dire celle qui a inspiré plusieurs des définitions plus

récentes (*e.g.*, Bloemer & Kasper, 1995; Fornell, 1992; Mano & Oliver, 1993; Oliver, 1997; Ostrom & Iacobucci, 1995) est (Hunt, 1977, p. 459) : « Une sorte de recul pris par rapport à l'expérience de consommation pour l'évaluer... l'évaluation que l'expérience était au moins aussi bonne qu'elle devait l'être », traduction libre de « *A kind of stepping away from an experience and evaluating it... the evaluation rendered that the experience was at least as good as it was supposed to be* ». Ainsi, comme pour la majorité des définitions, celle de Hunt conceptualise la satisfaction des clients comme un jugement positif ou négatif au regard de ce que ceux-ci ont reçu. Selon les auteurs qui adhèrent à cette définition, le degré auquel l'individu obtient ce qu'il voulait (ou davantage que ce qu'il voulait) est directement lié à son degré de satisfaction.

Aussi, le plus souvent, les auteurs considèrent la satisfaction et l'insatisfaction comme étant les opposés d'un même continuum (Westbrook & Oliver, 1991). Toutefois, s'inspirant de la théorie des deux facteurs de Herzberg (1966), laquelle a été développée dans le contexte des travaux sur la satisfaction envers son emploi, quelques auteurs ont considéré la satisfaction et l'insatisfaction comme étant deux construits différents et indépendants influencés par des éléments distincts (Vanhamme, 2002). Les études conduites pour tester cette hypothèse n'ont pas permis de conclure à la supériorité de la théorie des deux facteurs sur la vision traditionnelle de la satisfaction des clients (Mackoy & Spreng, 1995; Maddox, 1981).

Enfin, dans son analyse des définitions de la satisfaction, Oliver (1997) fait ressortir la notion de point de vue vertical de la satisfaction. Ce dernier « renvoie au degré d'agrégation des expériences de consommation/d'achat » (Vanhamme, 2002). Autrement dit, les définitions de la satisfaction des clients ne concernent pas toutes la satisfaction envers une situation spécifique (*e.g.*, Churchill & Surprenant, 1982; Halstead *et al.*, 1994), certains chercheurs définissant la satisfaction par rapport à une accumulation d'expériences avec un produit/service (*e.g.*, Anderson *et al.*, 1994). Les indices nationaux de satisfaction des clients envers diverses institutions et industries sont des exemples de ce genre d'agrégation.

En résumé, la littérature conçoit généralement la satisfaction des clients comme le jugement positif ou négatif au regard de ce qu'ils ont reçu. Quant à l'insatisfaction et la satisfaction, elles peuvent être vues comme étant les jugements antinomiques du même continuum.

L'opérationnalisation et la mesure du concept de satisfaction des clients

Il existe deux approches à l'opérationnalisation du concept de satisfaction des clients (Czepiel & Rosenberg, 1977; Vanhamme, 2002), soit l'approche globale et l'approche multi-items. Les prochains paragraphes traitent de ces deux stratégies d'opérationnalisation.

L'opérationnalisation selon l'approche globale

Cette première approche d'opérationnalisation du concept de satisfaction du client suggère de mesurer la satisfaction des clients de manière globale, c'est-à-dire envers le produit ou le service dans son ensemble. Les attributs du produit ou du service sont considérés comme indissociables et structurés de manière à former un tout que seule une évaluation globale peut arriver à saisir.

La majorité des mesures issues de l'approche globale sont composées d'un seul item. La mesure la plus courante étant celle où l'on demande aux clients d'indiquer leur degré de satisfaction sur une échelle de type Likert allant de « Très insatisfait » à « Très satisfait » (Westbrook, 1980; Westbrook & Oliver, 1981). Il existe cependant plusieurs variations de cette échelle, par exemple l'échelle en pourcentage où 100 % indique « Complètement satisfait » et 0 % « Pas du tout satisfait ».

Toutefois, utiliser un seul item plutôt que plusieurs pour mesurer la satisfaction des clients pose potentiellement un problème au plan psychométrique. Selon Fowler (1995), les mesures multi-items peuvent davantage couvrir l'ensemble du concept à mesurer que les mesures à un seul item (meilleure validité). Ensuite, puisque la mesure du concept se fait à partir de plusieurs indicateurs, il indique qu'il est potentiellement possible de diminuer l'erreur de mesure aléatoire (meilleure fidélité). Bref, la mesure multi-items permet éventuellement d'obtenir de meilleurs indices de validité et de fidélité. Toutefois, malgré ces critiques, très peu d'échelles multi-items ont été développées dans l'approche globale; celles de Oliver (1980, 1997) (voir le tableau A2.1 de l'annexe 2) et de Westbrook et Oliver (1981) sont les plus connues. Une des raisons à ceci est sans doute la difficulté de développer plusieurs items mesurant la satisfaction des clients de manière globale, autres que « Globalement, quel est votre degré de satisfaction envers ... ? ». En effet, les items deviennent rapidement répétitifs ou empiètent sur d'autres concepts proches, tels que l'évaluation de la qualité du produit ou service ou encore l'intention de rachat.

L'opérationnalisation selon l'approche multi-items

La deuxième stratégie d'opérationnalisation du concept de satisfaction du client, l'approche multi-items, est la plus répandue dans la littérature scientifique (Hazelrigg & Hard, 2000) de même que dans les entreprises (Vanhamme, 2002). Celle-ci consiste à décomposer le produit ou service en ses divers attributs. Le degré de satisfaction des clients est ensuite obtenu par la combinaison des degrés de satisfaction envers chacun de ces attributs (Churchill & Surprenant, 1982, p. 493). Par exemple, un produit tel qu'un téléviseur est décomposé en ses différents attributs comme la qualité de l'image, son esthétisme, sa facilité d'utilisation, etc. L'individu indique ensuite son degré de satisfaction envers chacun de ces attributs. Tout comme dans l'approche globale, le format de réponse peut varier beaucoup d'une étude à l'autre. Le format le plus courant demeure celui associé à l'échelle de type Likert dont la sémantique des points de réponse est associée au terme « satisfaction »; par exemple, une échelle allant de « Très insatisfait » à « Très satisfait ».

Le score de l'individu est ensuite obtenu en calculant la somme ou la moyenne de ses évaluations des divers attributs (Vanhamme, 2002). Moins fréquemment, il arrive qu'il soit aussi demandé aux répondants d'indiquer l'importance qu'ils accordent aux différents attributs. Le score de satisfaction de l'individu est alors calculé par la somme ou la moyenne pondérée, ceci afin de tenir compte du fait que les différents attributs d'un produit ou d'un service n'ont pas tous la même importance aux yeux des clients. Par contre, l'utilisation d'une pondération a été critiquée. Chu (2002) souligne qu'elle introduit un biais humain, car les clients ont tendance à indiquer une importance similaire pour l'ensemble des attributs. Il ajoute aussi que le fait de demander deux réponses complique la tâche des répondants. En conséquence, ceci augmente le temps de réponse et diminue le taux de participation. Qui plus est, cette approche n'améliore que très peu la validité prédictive de la mesure de satisfaction des clients (Yuksel & Rimmington, 1998).

Les mesures de la satisfaction des clients opérationnalisées selon l'approche multi-items présentent des avantages par rapport aux mesures issues de l'approche globale. Tout d'abord, elles s'avèrent très pertinentes en termes d'outils diagnostiques (Ironson, Smith, Brannick, Gibson, & Paul, 1989). En d'autres mots, elles permettent d'identifier les attributs jugés les plus ou les moins satisfaisants par les clients. Par la suite, elles comptent davantage d'items. Ce faisant, ces mesures peuvent potentiellement discriminer plus finement entre les sujets par rapport à leur degré de satisfaction (Bond & Fox, 2007). Conséquemment, plusieurs auteurs recommandent d'utiliser (Chakrapani, 1998; Hayes, 1998; Hill & Alexander, 2000; Vavra, 1997) et utilisent eux-mêmes (*e.g.*, Busacca & Padula, 2005; Chu, 2002; Danaher, 1997; Danaher & Haddrell, 1996; Page & Spreng, 2002; Yuksel & Rimmington, 1998) ce type de mesure de la satisfaction des clients.

Les critiques des mesures de la satisfaction des clients issues de l'approche multi-items concernent le choix des attributs évalués et le fait que la combinaison linéaire des scores aux attributs serait trop simpliste pour représenter adéquatement la réalité. Par exemple, Diener *et al.* (1985), Ironson, Smith, Brannick, Gibson et Paul (1989) de même que

Scarpello et Campbell (1983) précisent que certaines dimensions pourraient avoir été omises dans l'échelle ou encore que certaines, bien qu'incluses, pourraient être sans grande importance aux yeux des clients. Ensuite, le nombre de dimensions utilisées et leur importance pourraient varier en fonction des individus, des ethnies, des segments de clientèle, etc. (Applebaum, Straker, & Geron, 2000; Diener *et al.*, 1985). Ces auteurs ajoutent qu'une simple combinaison linéaire des dimensions risque de ne pas être représentative de la stratégie cognitive utilisée par chaque individu pour combiner ces dimensions et exprimer son jugement global. Il est probable qu'un ensemble de relations compensatoires (p. ex. un faible degré de satisfaction sur un attribut peut être compensé par un degré élevé de satisfaction sur un autre attribut) ou non compensatoires (p. ex. un faible degré de satisfaction sur un attribut ne peut être compensé par un degré de satisfaction élevé sur un autre attribut) entrent en jeu (Hennig-Thurau & Klee, 1997). Il n'est donc pas rare qu'une mesure globale composée d'un ou de quelques items soit jumelée à une mesure multi-items afin de pallier à ces problèmes.

Le problème de recherche

Une des principales préoccupations soulevées dans la littérature au sujet de la mesure de la satisfaction des clients est que les scores de satisfaction sont généralement élevés. Autrement dit, les résultats des études de satisfaction des clients présentent à peu près tous une asymétrie négative par rapport à la distribution normale, et ce, peu importe l'objet sur lequel porte le jugement de satisfaction (Danaher & Haddrell, 1996; Larsen, Attkisson, Hargreaves, & Nguyen, 1979; Nguyen, Attkisson, & Stegner, 1983; Peterson & Wilson, 1992; Vanhamme, 2002; Westbrook & Oliver, 1981; Yi, 1990). Tel qu'illustré à la figure 1.1, en général, les réponses d'insatisfaction sont peu nombreuses et la moyenne des scores de satisfaction des clients est, par conséquent, toujours élevée. Globalement, les moyennes se situent à environ 75 % du maximum de l'échelle de satisfaction (Mellor, Cummins, & Loquet, 1999), des taux moyens de satisfaction de 80 % et plus étant chose courante (Worthington, 2005). Les données de l'ACSI, ou *American Customer Satisfaction Index*,

illustrent bien ces constats. À titre d'exemple, les résultats publiés pour le troisième trimestre de l'année 2009 indiquent que le taux moyen de satisfaction des Américains envers l'ensemble des biens et services de consommation se situait à 76,0 %. En fait, comme on peut le voir à la figure 1.2, ce taux moyen s'est toujours situé entre 70,0 % et 76,1 % depuis l'instauration de l'ACSI en 1994 (ACSI, 2009).

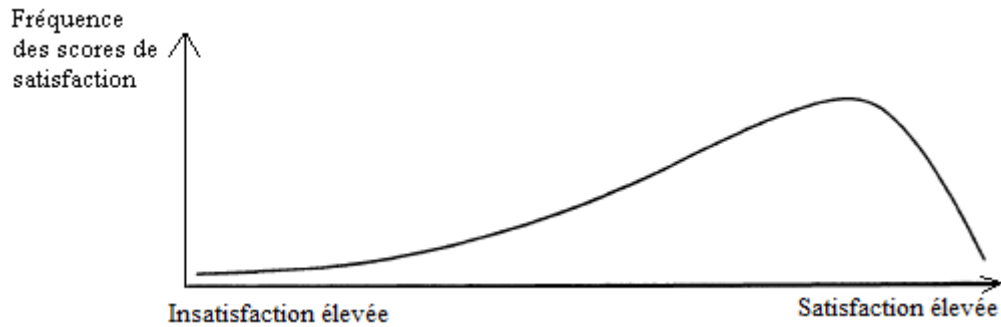


Figure 1.1 La distribution dissymétrique négative des scores à un questionnaire de satisfaction des clients (adaptée de Peterson & Wilson, 1992)

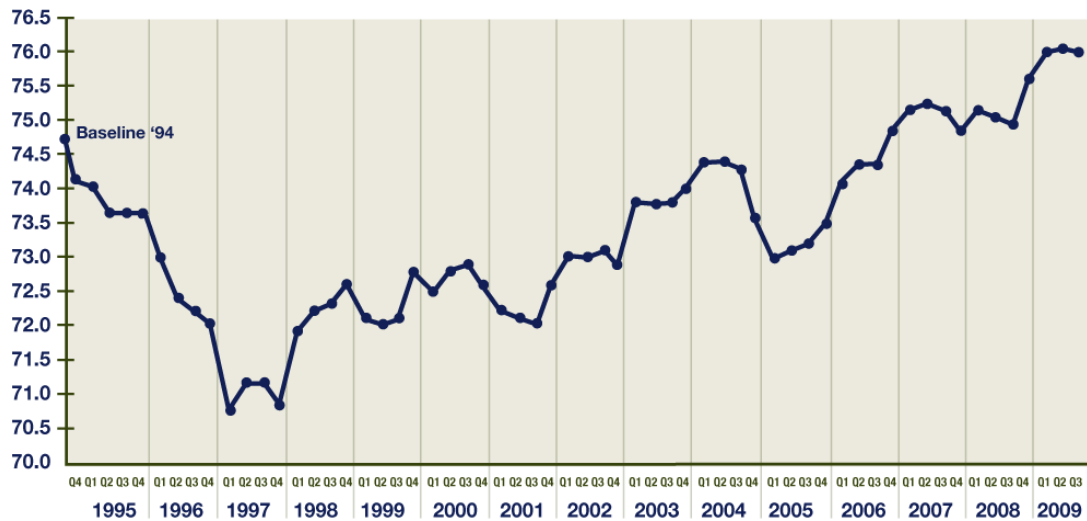


Figure 1.2 L'évolution de l'*American Customer Satisfaction Index* pour la période 1994-2009 (tirée du site Internet de l'ACSI)

Ce phénomène, loin d'être nouveau, a été rapporté dès les premières études de satisfaction des clients (*e.g.*, Frank, Salzman, & Fergus, 1977; Larsen *et al.*, 1979; Linn, 1975). Il a même été qualifié de caractéristique frappante par Peterson et Wilson (1992). Les principales hypothèses pour expliquer ce phénomène sont le biais de positivité, le biais d'échantillonnage et le fait que la satisfaction est naturellement plus fréquente que l'insatisfaction.

Premièrement, selon le biais de positivité, aussi appelé biais d'indulgence (Tourangeau, Rips, & Rasinski, 2000) ou biais de clémence (Guilford, 1954), les individus sont naturellement peu enclins à donner des évaluations négatives aux autres. À titre d'exemple, ce biais est souvent rapporté en évaluation du personnel (French-Lazovik & Gibson, 1984). Ainsi d'après le biais de positivité, les individus auraient tendance à biaiser à la hausse leurs réponses dans les études de satisfaction.

Deuxièmement, la distribution non normale des scores de satisfaction pourrait être due à un biais d'échantillonnage (Larsen *et al.*, 1979). L'explication avancée est que les clients insatisfaits cesseraient rapidement d'utiliser les services d'une organisation ou d'acheter un produit. Ainsi, la clientèle échantillonnée serait principalement composée de clients assez satisfaits. Le fait que la plupart des organisations n'incluent pas dans leurs études de satisfaction les clients qui ont cessé d'acheter leurs produits ou leurs services (Horovitz, 2000) renforce cette idée. En effet, il est habituellement très difficile et coûteux, voire parfois impossible, d'identifier les clients perdus et de les retracer pour les inclure dans l'échantillon. Il est alors probable que les degrés élevés de satisfaction observés dans les diverses études soient dus au fait que les clients perdus, et donc les plus susceptibles d'être insatisfaits, sont habituellement exclus des échantillons.

Enfin, il apparaît que la satisfaction pourrait être naturellement plus fréquente que l'insatisfaction (Peterson & Wilson, 1992). En d'autres termes, la distribution de la

satisfaction exprimée par des clients ne suivrait pas la courbe normale, contrairement à ce que l'on observe avec plusieurs autres construits psychologiques.

Cette hypothèse trouve appui en économie et en psychologie. D'abord, d'un point de vue macroéconomique, Fornell (1992) et Gould (1995) avancent que dans les marchés où des offres concurrentes sont disponibles, il est logique que les scores de satisfaction aient une distribution dissymétrique négative. Ces auteurs rapportent que, dans ces conditions, les clients insatisfaits de l'offre ou du service d'une entreprise sont tentés d'aller voir ce qu'ont à offrir les concurrents. Ainsi, les entreprises qui ne réussissent pas à satisfaire leurs clients ne peuvent pas survivre puisqu'elles n'auront pas assez de clients pour être rentables. Autrement dit, être capable de satisfaire les clients serait le minimum requis pour avoir une chance de survivre dans un marché où il y a de la concurrence. En fait, même en situation de monopole, l'insatisfaction des clients peut mener à la fermeture de l'entreprise. C'est le cas lorsqu'il existe une alternative vers laquelle les clients insatisfaits peuvent se tourner, par exemple un bien substitut. Ensuite, selon l'un des postulats de base du modèle microéconomique, le ménage tente toujours de maximiser sa satisfaction en tenant compte de ses ressources et de ses contraintes (Bryant, 1990). En d'autres termes, avant de choisir un produit ou un service, les individus estiment la satisfaction que leur apporterait les diverses options possibles. Ils choisissent ensuite l'option pour laquelle la satisfaction attendue est, selon eux, la plus élevée. Selon cette approche, les situations de satisfaction sont nécessairement plus fréquentes que les situations d'insatisfaction, l'insatisfaction étant exactement ce que les individus tentent de minimiser.

Les chercheurs qui s'intéressent à ce phénomène sous l'angle de la psychologie arrivent aux mêmes constats. Ils avancent que la satisfaction et l'insatisfaction sont respectivement considérées comme des états plaisants et déplaisants (Oliver, 1997; Watson & Tellegen, 1985). Tel qu'on pourrait le penser intuitivement et comme le propose le principe du plaisir en psychanalyse (Freud, Strachey, & Richards, 1974), les individus cherchent le plaisir et tentent d'éviter le déplaisir. Au bout du compte, les situations satisfaisantes et donc

plaisantes sont sûrement plus fréquentes, car elles sont recherchées par les individus. À l’opposé, les situations insatisfaisantes ne sont pas recherchées par les individus. Bref, la perspective psychologique rejoint la perspective microéconomique en postulant, elle aussi, que les gens choisissent les options qui ont la plus forte probabilité de maximiser leur satisfaction.

Cette forte concentration de scores positifs est considérée problématique généralement pour trois raisons. Premièrement, elle peut conduire à une mauvaise interprétation des scores. Effectivement, un score élevé de satisfaction des clients ne peut pas être systématiquement interprété comme une preuve que l’organisation performe bien (Larsen *et al.*, 1979; Worthington, 2005). Deuxièmement, selon l’ampleur de cette asymétrie par rapport à la courbe normale, il devient parfois plus compliqué d’utiliser les tests statistiques paramétriques.

La troisième raison pour laquelle cette forte concentration de scores positifs est considérée problématique est de nature psychométrique : le fait que les scores sont élevés, c’est-à-dire que les individus utilisent majoritairement la partie positive de l’échelle de réponse, restreint la variance des items et du questionnaire dans son ensemble (Giese & Cote, 2000; Yi, 1990). Effectivement, la variance d’un item ne peut être maximisée que lorsque sa moyenne se situe au centre de l’échelle de réponse (DeVellis, 1991; Weisberg, 1992). Lorsque c’est le cas, on dit que l’item est de difficulté intermédiaire. Courville (2004) en fait la démonstration pour le cas le plus simple, soit celui d’un item dichotomique (tableau 1.1). Pour ce type d’item, la variance (s^2) est égale à $p(1-p)$, où p correspond à la moyenne de l’item, c’est-à-dire la proportion des réponses qui se retrouvent dans la première (ou la deuxième, selon le cas) des deux catégories de réponse. Tel qu’il est montré au tableau 1.1, la variance est maximisée lorsque la moyenne se situe à 0,50, soit quand la moitié des répondants ont un score de 0 et l’autre moitié un score de 1. Il faut cependant préciser que l’écart entre le score moyen d’un item et son score central influence moins l’ampleur de la variance dans le cas où cet item est polytomique que dans celui où il est

dichotomique (Nunnally, 1978). Selon Peterson (1997), dans le cas d'items polytomiques, la corrélation entre l'écart-type des items, d'une part, et la différence observée entre les moyennes d'items et le score central de l'échelle de réponse est égale à $-0,30$ ($p < 0,01$). Cette relation a aussi été rapportée par Cummins (2003) suite à l'analyse des résultats de plusieurs études de satisfaction. Ce dernier a remarqué que plus la moyenne des scores de satisfaction (exprimés en pourcentage) se rapproche du score central de la mesure de satisfaction, plus l'écart-type était important (voir la figure 1.3).

Tableau 1.1 Lien entre la moyenne et la variance pour un item dichotomique (tiré de Courville, 2004)

Moyenne de l'item	Variance de l'item
p	s^2
0,00	0,00
0,10	0,09
0,20	0,16
0,30	0,21
0,40	0,24
0,50	0,25
0,60	0,24
0,70	0,21
0,80	0,16
0,90	0,09
1,00	0,00

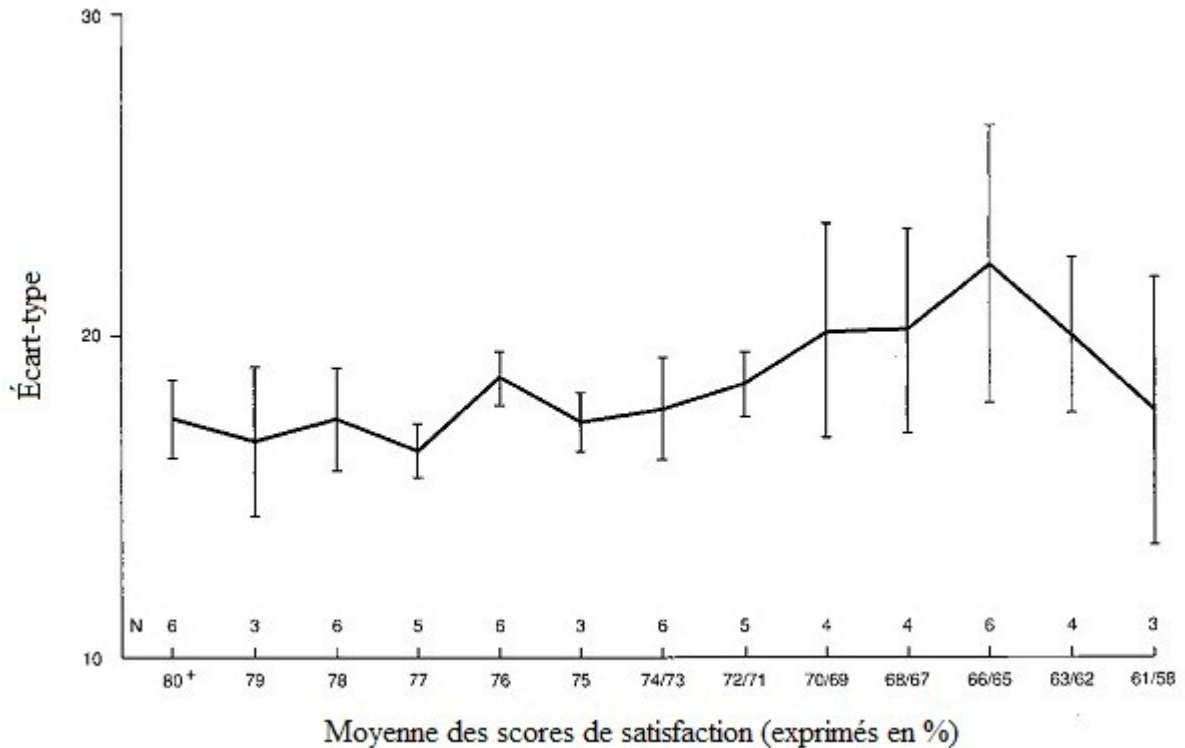


Figure 1.3 Relation entre la moyenne des scores de satisfaction (exprimés en pourcentage) et l'ampleur de l'écart-type (tirée de Cummins, 2003)

De même, plusieurs études ont démontré que la réduction de la variance causée par la distribution asymétrique d'une variable entraîne une sous-estimation des covariances et des coefficients de corrélation (Dunlap, Burke, & Greer, 1995; Kowalski, 1972; Schwarz, 1996). Les indices psychométriques étant basés sur la covariance et sur la corrélation (p. ex. la cohérence interne, la stabilité temporelle, la validité prédictive, la validité concomitante), ils sont eux aussi affectés à la baisse par cette restriction de la variance.

À titre d'exemple, plusieurs auteurs rapportent que le coefficient alpha de Cronbach augmente de façon non linéaire avec l'importance des variances d'items et des corrélations inter-items ainsi qu'avec le nombre d'items (Armstrong, Jones, & Wang, 1998; Bruner & Hensel, 1993; Cortina, 1993; Cox, 1980; Cronbach, 1951; DeVellis, 1991). Dans cette

même veine, Greer *et al.* (2006) ont trouvé qu'une distribution dissymétrique des items affecte à la baisse les corrélations inter-items réduisant en retour l'ampleur du coefficient alpha de Cronbach. Plus précisément, cette relation était d'autant plus forte que la dissymétrie était importante, que le nombre d'items était petit, que les corrélations inter-items étaient faibles et que les items possédaient des asymétries allant dans des directions opposées. Toujours dans cet ordre d'idées, Berg *et al.* (2007) spécifient que, de manière générale, une plus grande dispersion des scores ainsi que des moyennes d'items situées au centre de l'échelle de réponse sont associées à une augmentation de la variance des scores et de la fidélité d'une mesure. Enfin, Yu (2003) explique que le coefficient alpha de Cronbach peut subir une diminution importante (p. ex. de 0,60 à 0,30) simplement à cause d'une réduction de la variance des items provoquée par l'utilisation d'un groupe de sujets plus homogènes relativement à la caractéristique mesurée.

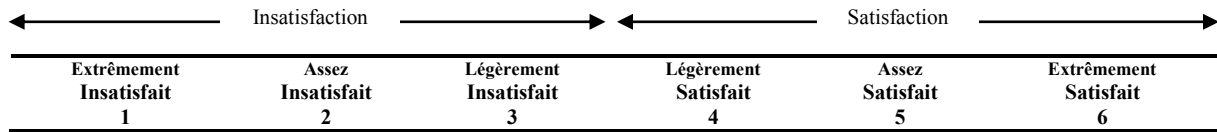
En somme, une restriction de la variance des items influence à la baisse l'estimation des qualités psychométriques d'un outil de mesure, c'est-à-dire qu'elle tend à faire augmenter l'estimation de l'erreur de mesure. Pour ces raisons, plusieurs auteurs s'intéressant à la satisfaction des clients ont avancé que la dissymétrie négative des scores réduit la capacité des outils de mesure à discriminer entre eux les sujets par rapport à leur degré de satisfaction (Ganglmair & Lawson, 2003; Nguyen *et al.*, 1983). De même, d'autres ont avancé que la dissymétrie négative des scores amène probablement une sous-estimation des corrélations entre la satisfaction des clients et les variables qui y sont associées (Fornell, 1992; Peterson & Wilson, 1992; Yi, 1990). En effet, ces mesures discriminent mal entre les clients qui seront fidèles et ceux qui ne le seront pas (Danaher & Haddrell, 1996). Comme Stauss et Neuhaus (1997) le mentionnent, malgré que dans la majorité des cas 80 % à 90 % des clients rapportent être « très satisfaits » ou « satisfaits », un nombre non négligeable d'entre eux cesse d'acheter chez l'entreprise ou change de marque.

Dans le but d'améliorer cette situation, plusieurs auteurs proposent d'augmenter la sensibilité des échelles de réponse des mesures de satisfaction afin d'augmenter la variance

associée à la caractéristique mesurée (*i.e.*, la variance vraie), la satisfaction. Pour ce faire, certains auteurs proposent de manipuler la polarité de l'échelle de réponse, alors que d'autres suggèrent d'en manipuler le nombre de points.

La proposition de manipuler la polarité de l'échelle de réponse émane de l'approche psychologique. Une échelle de réponse peut être bipolaire - elle couvre alors les parties positives et négatives du jugement - ou unipolaire - elle couvre alors uniquement la partie positive ou négative du jugement (Schaeffer & Presser, 2003). Les échelles de réponse bipolaires sont celles qui sont les plus souvent utilisées dans les questionnaires de satisfaction des clients. Dans ce contexte d'utilisation, une échelle de réponse bipolaire couvre l'insatisfaction et la satisfaction (voir la figure 1.4 à titre d'exemple). La première moitié de l'échelle de réponse représente l'insatisfaction et la seconde moitié la satisfaction. Une échelle de réponse unipolaire se concentre soit sur l'insatisfaction, soit sur la satisfaction. La figure 1.4 donne un exemple du cas le plus courant, soit celui d'une échelle de réponse unipolaire qui comprend uniquement des choix de réponse se rapportant à la satisfaction. Comme il peut être remarqué dans cet exemple, les divers degrés d'insatisfaction sont regroupés en un seul point, soit « Aucunement satisfait », qui exprime l'absence de satisfaction. Quant aux autres points, ils représentent des degrés faibles à élevés de satisfaction. L'échelle de réponse unipolaire porte donc sur une moitié du continuum évaluatif contrairement à l'échelle bipolaire qui possède le même nombre de choix de réponse pour les deux moitiés de ce continuum. À cause de cette caractéristique de l'échelle de réponse unipolaire, certains auteurs pensent qu'elle permet une meilleure discrimination entre les sujets sur une des deux moitiés du continuum évaluatif (French-Lazovik & Gibson, 1984; Gannon & Ostrom, 1996; Lam & Klockars, 1982; Weiss, Dawis, England, & Lofquist, 1967). Suivant cette logique, ces auteurs suggèrent l'utilisation d'une échelle de réponse unipolaire se concentrant sur la partie positive du continuum évaluatif lorsque la majorité des individus évalués s'y retrouvent.

Exemple d'échelle de réponse bipolaire :



Exemple d'échelle de réponse unipolaire :

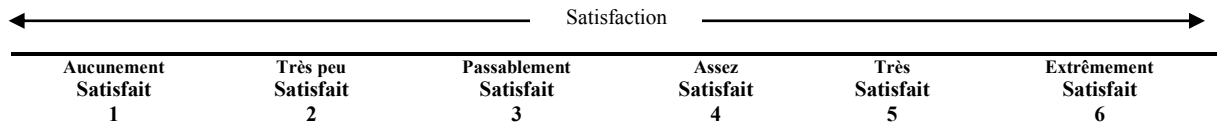


Figure 1.4 Illustration de la première proposition : utiliser une échelle de réponse unipolaire plutôt que bipolaire

La proposition de manipuler le nombre de points de l'échelle de réponse émane quant à elle de l'approche économique. Plus précisément, selon les auteurs qui la soutiennent, la sensibilité de l'échelle de réponse d'une mesure de satisfaction peut être améliorée en augmentant son nombre de points (Cummins, 2002, 2003; Cummins & Gullone, 2000; Fornell, 1992; Fornell, Johnson, Anderson, Cha, & Bryant, 1996). La figure 1.5 illustre cette idée. Le premier argument évoqué en faveur de cette suggestion est le suivant : augmenter le nombre de points dans l'échelle de réponse donne au répondant davantage de latitude pour préciser son degré de satisfaction et, conséquemment, accroît la proportion de variance vraie et le degré de discrimination des items. Le deuxième argument consiste à ce que la catégorisation d'une variable continue implique nécessairement une perte d'information, c'est-à-dire une perte de variance vraie. En effet, moins il y a de catégories, moins il y a d'information transmise. Suivant cette logique, il semble qu'augmenter le nombre de points de réponse permettrait de se rapprocher de la quantité d'information transmise par une variable continue (Cohen, 1983; Shaw, Huffman, & Haviland, 1987). Bref, l'idée qui sous-tend ces deux arguments est qu'une augmentation du nombre de points de réponse permettrait l'obtention d'une mesure plus sensible aux différences individuelles. En conséquence, ceci se traduirait par une augmentation de la variance vraie et des qualités psychométriques de l'instrument de mesure. S'appuyant sur cette logique,

bon nombre de chercheurs (e.g., Mazaheri & Theuns, 2005; Mazaheri & Theuns, 2009; McKinney, Yoon, & Zahedi, 2002; Oterhals, Hanestad, Eide, & Hanssen, 2006) ont opté pour des échelles de réponse en dix ou onze points pour mesurer la satisfaction. C'est le cas des concepteurs des indices nationaux de satisfaction des clients, comme l'*American Customer Satisfaction Index* ou le *Swedish Customer Satisfaction Index*, et des chercheurs affiliés à des firmes privées de recherche telles que la firme multinationale *The Leadership Factor*.

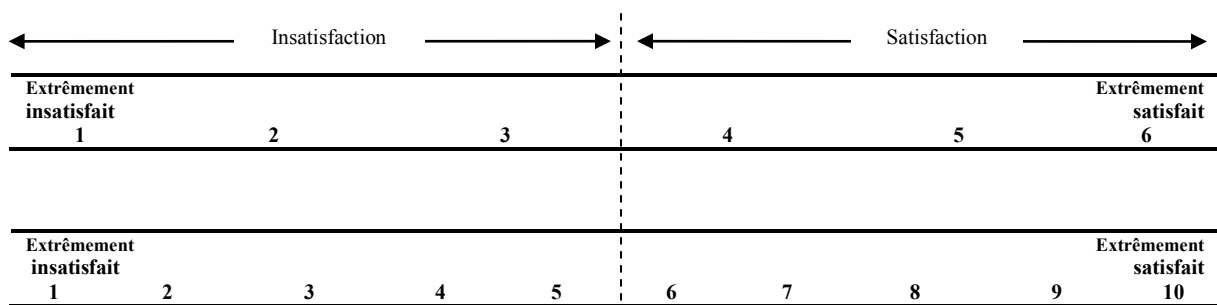


Figure 1.5 Illustration de la deuxième proposition : augmenter le nombre de points dans l'échelle de réponse

Ces deux suggestions trouvent aussi un appui théorique dans le cadre de l'approche cognitive des méthodes de sondage. Cette dernière, qui s'intéresse, entre autres, à la façon dont les individus génèrent leurs réponses aux questions dans une enquête par questionnaire, a permis de démontrer que le format de réponse utilisé peut influencer les paramètres de la distribution des réponses (p. ex. la moyenne, la variance) de même que la précision des réponses. Toutefois, malgré le fait que ces deux propositions apparaissent logiques théoriquement et qu'elles soient supportées par quelques études, leur impact réel sur la distribution des scores de satisfaction des clients, de même que sur les qualités psychométriques d'une mesure de satisfaction des clients, demeure peu connu. De surcroît, l'effet des diverses combinaisons possibles de ces deux variables (la polarité et le nombre de points de l'échelle de réponse) sur les qualités psychométriques d'une mesure de satisfaction des clients n'a pas été évalué de manière systématique. Il apparaît donc justifié

de nous demander si la polarité et le nombre de points dans l'échelle de réponse ont un effet observable sur les paramètres de distribution des scores de satisfaction des clients et sur les qualités psychométriques d'une mesure de satisfaction des clients.

Objectif général et pertinence de l'étude

L'objectif général de cette étude est d'évaluer si la polarité et le nombre de points d'une échelle de réponse ont une influence sur les paramètres de distribution des réponses et sur les qualités psychométriques d'une mesure de la satisfaction des clients. Afin d'atteindre cet objectif, cette étude prend appui sur l'approche cognitive des méthodes de sondage, ce courant d'études étant de loin celui qui s'est le plus intéressé aux effets de l'échelle de réponse sur les réponses obtenues.

D'un point de vue scientifique, cette étude permettra de mettre à l'épreuve les deux propositions faites dans la littérature pour augmenter la variance dans un questionnaire de satisfaction des clients et d'en améliorer les qualités psychométriques. De plus, contrairement aux recherches précédentes qui ne considéraient que la polarité ou le nombre de points de l'échelle de réponse, la présente éprouvera les différentes combinaisons possibles de ces deux facteurs. Dans le cas où une relation serait identifiée entre ces caractéristiques de l'échelle de réponse et les qualités psychométriques de la mesure, il sera alors possible d'identifier l'échelle de réponse à privilégier pour maximiser les qualités psychométriques et donc minimiser l'erreur de mesure. Dans le cas où l'étude constaterait une absence de relation entre les caractéristiques de l'échelle de réponse et les qualités psychométriques de la mesure, ceci ferait avancer le débat sur la mesure de la satisfaction des clients, les futures recherches visant à améliorer les mesures de la satisfaction des clients devant alors se tourner vers d'autres avenues.

D'un point de vue pratique, cette étude va aider les concepteurs d'études de satisfaction dans leurs choix méthodologiques. Ils auront en effet une meilleure idée de l'impact que peut avoir ou non le choix d'une échelle de réponse sur la qualité des données recueillies à l'aide d'un questionnaire de satisfaction des clients.

Chapitre 2 : Cadre théorique

Dans ce deuxième chapitre, nous présentons le cadre théorique de cette étude. Il débute avec la présentation du modèle théorique de la satisfaction des clients sur lequel s'appuie l'étude. Ce modèle permettra d'identifier les variables qui serviront à évaluer la validité nomologique de la mesure de satisfaction des clients. L'approche cognitive des méthodes de sondage est ensuite abordée. Cette section débute par une présentation du modèle de Tourangeau, Rips et Rasinski (2000) décrivant le processus de réponse à un sondage. Elle se poursuit avec une analyse des études qui se sont intéressées à l'influence de la polarité de l'échelle de réponse sur la distribution des réponses et celles qui ont évalué la relation entre le nombre de points de l'échelle de réponse et les qualités psychométriques d'un outil de mesure. Ce chapitre se termine avec la présentation de l'objectif spécifique de l'étude et des hypothèses de recherche.

Le modèle théorique de la satisfaction des clients

Le modèle théorique de la satisfaction des clients de Oliver (1980, 1997) sera utilisé dans cette étude (voir la figure 2.1). Le choix de ce modèle repose sur deux raisons. Tout d'abord, il a été mis à l'épreuve plusieurs fois par les chercheurs, ce qui constitue, à nos yeux, une raison suffisante pour le retenir, les comparaisons avec les autres études étant alors possibles (Athiyaman, 1997; Bearden & Teel, 1983; Hallowell, 1996; LaBarbera & Mazursky, 1983; Oliver & Bearden, 1985; Söderlund & Vilgon, 1999). Deuxièmement, ce modèle prend appui sur une théorie psychosociale reconnue, soit la théorie de l'action raisonnée (Fishbein & Ajzen, 1975). Bon nombre d'études ont démontré au fil des années la capacité de cette théorie à prédire les comportements humains, dont ceux des consommateurs (*e.g.*, Ha, 1998; Oliver & Bearden, 1985).

Les prémisses du modèle d'Oliver sont les mêmes que ceux de la théorie de l'action raisonnée. La première prémisse consiste en ce que les individus sont généralement rationnels, qu'ils utilisent l'information qui les entoure pour prendre des décisions et qu'ils soupèsent habituellement les conséquences positives et négatives associées à l'adoption éventuelle d'un comportement avant de décider s'ils passent ou non à l'action. La deuxième prémisse suppose que le meilleur prédicteur d'un comportement d'ordre volitif est l'intention d'effectuer ce comportement. Oliver (1997) note que dans certains cas d'autres variables peuvent agir sur l'intention. Il cite l'influence possible de la perception des normes sociales, une autre variable de la théorie de l'action raisonnée. Oliver émet cependant l'hypothèse que la relation entre la satisfaction des clients et leur comportement de rachat s'explique essentiellement par l'attitude envers le comportement et l'intention comportementale. Les résultats d'une étude que nous avons réalisée vont dans ce sens (Renaud, 2005). Les prochains paragraphes présentent le modèle d'Oliver et les variables qui le composent, c'est-à-dire la satisfaction du client, l'attitude envers le comportement, l'intention comportementale et le comportement.

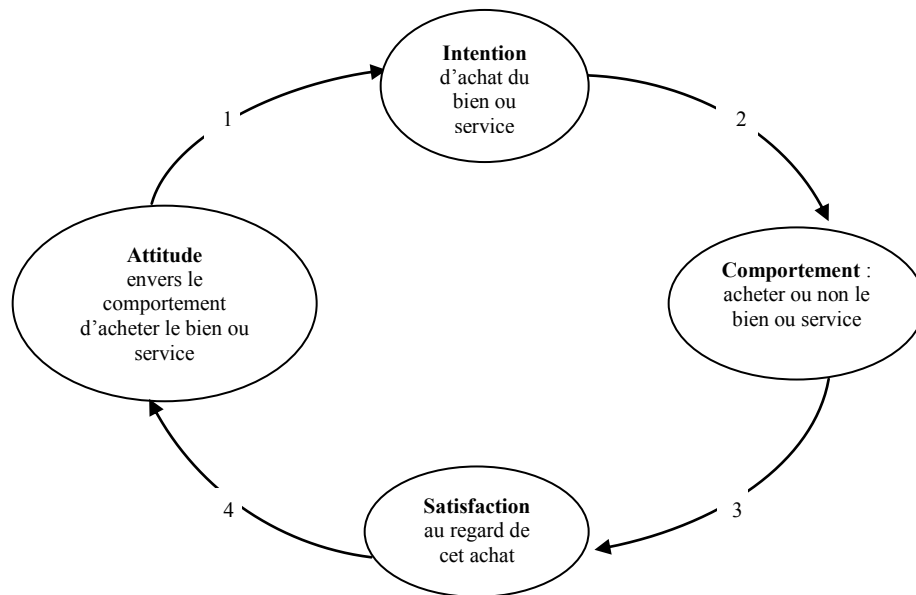


Figure 2.1 Représentation circulaire (dynamique) du modèle théorique de la satisfaction des clients d'Oliver (1980, 1997)

Le modèle d'Oliver se caractérise par un enchaînement de cycles d'achats. Au début du processus, le client forme son attitude envers le comportement d'acheter un bien ou un service. Cette attitude influence son intention d'achat (lien 1). Si cette intention est assez forte, le client décidera alors d'acheter le bien ou le service en question (lien 2). À la suite de cette expérience d'achat, le client ressentira ou exprimera un degré plus ou moins élevé de satisfaction au regard de cet achat (lien 3) qui l'amènera à maintenir ou à modifier son attitude envers l'achat de ce bien ou service (lien 4). Autrement dit, avant le tout premier cycle d'achat, le client possède une attitude envers le comportement d'acheter un bien ou un service. Cette attitude est basée sur diverses sources d'information (p. ex. la publicité, le bouche-à-oreille), dont l'expérience personnelle lorsque le bien ou service a déjà été acheté auparavant par l'individu (Oliver, 1997). L'attitude est définie comme une évaluation plus ou moins favorable ou défavorable relativement durable et stable dans le temps (Eagly & Chaiken, 1993). Plus spécifiquement, l'attitude envers le comportement de rachat correspond à la perception des avantages et des désavantages d'acheter le bien ou le service. Plus les avantages l'emportent sur les désavantages, plus l'attitude sera favorable. Telle que définie ici, l'attitude envers le comportement d'achat rejoint les concepts d'attentes du client en marketing (Athiyaman, 1997) et d'utilité attendue en économique (Rust, Inman, Jia, & Zahorik, 1999).

Comme illustré à la figure 2.1, l'attitude envers le comportement d'achat influence directement l'intention d'achat (lien 1). Plus l'attitude envers l'achat du bien ou du service est favorable, plus l'intention de l'acheter sera forte. À l'inverse, plus l'attitude envers l'achat du bien ou du service est défavorable, plus l'intention de l'acheter sera faible. L'intention comportementale est définie comme la probabilité subjective que l'individu effectue le comportement en question (Fishbein & Ajzen, 1975).

Lorsque l'intention d'effectuer un premier achat ou un rachat est assez forte, elle augmente la probabilité que le comportement d'achat s'actualise (voir le deuxième lien de la figure 2.1). Ensuite, comme le montre le troisième lien de la figure 2.1, plus le comportement

d'achat résulte en des bénéfices égaux ou supérieurs à ceux préalablement escomptés, plus l'individu sera satisfait (c.-à-d. qu'il fera une évaluation favorable du bien ou du service reçu). L'inverse est aussi vrai. Plus le comportement d'achat résulte en des bénéfices inférieurs à ceux escomptés, plus l'individu sera insatisfait (c.-à-d. qu'il fera une évaluation défavorable du bien ou du service reçu). La satisfaction du client, rappelons-le, est le plus souvent définie dans la littérature comme le jugement positif ou négatif au regard du bien ou du service reçu. Il s'agit donc d'un concept évaluatif comme l'attitude. Contrairement à l'attitude toutefois, la satisfaction nécessite l'expérience directe du bien ou du service. À titre d'exemple, quelqu'un peut avoir une attitude favorable envers le comportement qui consiste à acheter un ordinateur portable d'une marque donnée sans jamais en avoir fait l'expérience. Il ne peut pas, par contre, être satisfait ou insatisfait d'un ordinateur portable de ladite marque s'il n'en a jamais fait l'expérience.

Le degré de satisfaction exprimé par une personne agit ensuite comme une information ayant un effet sur son attitude envers le bien ou le service acheté (lien 4). Toutefois, plus son attitude est cristallisée, moins les nouvelles expériences d'achat (c.-à-d. les jugements de satisfaction subséquents) auront une influence sur son attitude. C'est donc dire que plus un individu possède une grande expérience d'achat à l'égard d'un bien ou d'un service, moins son degré de satisfaction vis-à-vis de cet achat influencera de manière importante son attitude (Rust *et al.*, 1999). Par exemple, supposons un premier individu qui a une attitude très favorable envers le fait de faire son épicerie chez le commerçant « A ». Supposons également que depuis environ cinq ans, cet individu va la plupart du temps faire son épicerie chez lui. Sa grande expérience de magasinage chez ce commerçant fait en sorte que son attitude très favorable envers le fait d'y faire son épicerie est maintenant durable et stable, c'est-à-dire cristallisée. Ainsi, qu'il en ressorte insatisfait lors d'une prochaine visite aurait peu d'impact sur son attitude. Considérons maintenant un deuxième individu qui n'est jamais allé faire son épicerie chez le commerçant « A », mais qui a une attitude assez favorable pour se laisser tenter d'y faire des achats. Si dès sa première visite chez lui il en ressort insatisfait, il est alors très probable que cette unique expérience aura un effet négatif important sur son attitude. L'impact pourrait même être assez important pour qu'il n'ait pas

l'intention d'y retourner, ce qui mettrait fin aux cycles d'achat. Pour ces deux individus, une seule expérience de magasinage insatisfaisante aura eu un impact fort différent sur leur attitude. Du point de vue du commerçant, ceci signifie que s'il réussit à satisfaire ceux-ci de manière constante, ses clients seront moins sensibles à une insatisfaction ponctuelle. Réussir à satisfaire ses clients la majorité du temps permet donc à une entreprise de conserver ou de fidéliser sa clientèle (Anderson & Sullivan, 1993).

Enfin, outre l'intention de rachat, l'intention de faire du bouche-à-oreille est aussi associée au degré de satisfaction. Le bouche-à-oreille réfère à la communication interpersonnelle effectuée par les clients (Solomon, Polegato, & Zaichkowsky, 1999). Il peut être soit positif, tel que recommander le bien ou le service ou en parler en bien à son entourage, soit négatif, tel que le déconseiller ou en parler en mal. La probabilité de faire du bouche-à-oreille positif augmente avec le degré de satisfaction. De même, la probabilité de faire du bouche-à-oreille négatif augmente avec l'insatisfaction (Söderlund, 1998). Le bouche-à-oreille est une variable qui tire son importance du fait que c'est un des facteurs qui influencent le plus la décision d'achat des individus (Council of Better Business Bureaus, 2001). C'est effectivement une des sources d'informations les plus crédibles pour les individus, et ce, principalement à cause de son caractère non marketing (Yi, 1990).

L'intention de faire du bouche-à-oreille ne fait pas explicitement partie du modèle d'Oliver parce que son modèle se concentre sur le lien entre le degré de satisfaction et la probabilité de rachat. C'est pour cette raison que l'intention de faire du bouche-à-oreille ne fait pas partie de la figure 2.1. Cependant, le bouche-à-oreille est identifié par Oliver (1997) ainsi que par la majorité des auteurs s'intéressant au sujet comme une variable associée à la satisfaction. Selon les auteurs, la relation entre le degré de satisfaction et l'intention de faire du bouche-à-oreille peut être directe ou indirecte. Toutefois, la plupart du temps, l'intention de faire bouche-à-oreille est associée directement à la satisfaction (*e.g.*, Szymanski & Henard, 2001; Wirtz & Lee, 2003; Yi, 1990). Les auteurs qui considèrent que cette relation est indirecte estiment que l'attitude agit comme une variable médiatrice entre la satisfaction

et l'intention de faire du bouche-à-oreille (Athiyaman, 1997). Notons que la plupart du temps, c'est l'intention de faire du bouche-à-oreille positif plutôt que négatif qui est pris en compte dans les études, c'est-à-dire l'intention de recommander le bien, le service ou l'entreprise. Ce choix s'explique probablement par le fait que les réponses d'insatisfaction sont peu fréquentes dans les études. Conséquemment, il est davantage pertinent de mesurer l'intention de faire du bouche-à-oreille positif que l'intention de faire du bouche-à-oreille négatif.

En somme, le modèle théorique de la satisfaction des clients utilisé dans cette étude est celui de Oliver (1980, 1997). Il contient quatre variables : l'attitude envers le comportement d'achat, l'intention d'achat, le comportement d'achat et la satisfaction envers l'achat. À ce modèle est ajoutée l'intention de faire du bouche-à-oreille positif étant donné que cette variable est fréquemment incluse dans le réseau nomologique du concept de satisfaction des clients. La figure 2.2 présente le modèle qui résulte de l'ajout de l'intention de faire du bouche-à-oreille positif au modèle d'Oliver. Les liens cinq et six représentent respectivement les relations directe et indirecte entre la satisfaction et l'intention de faire du bouche-à-oreille positif.

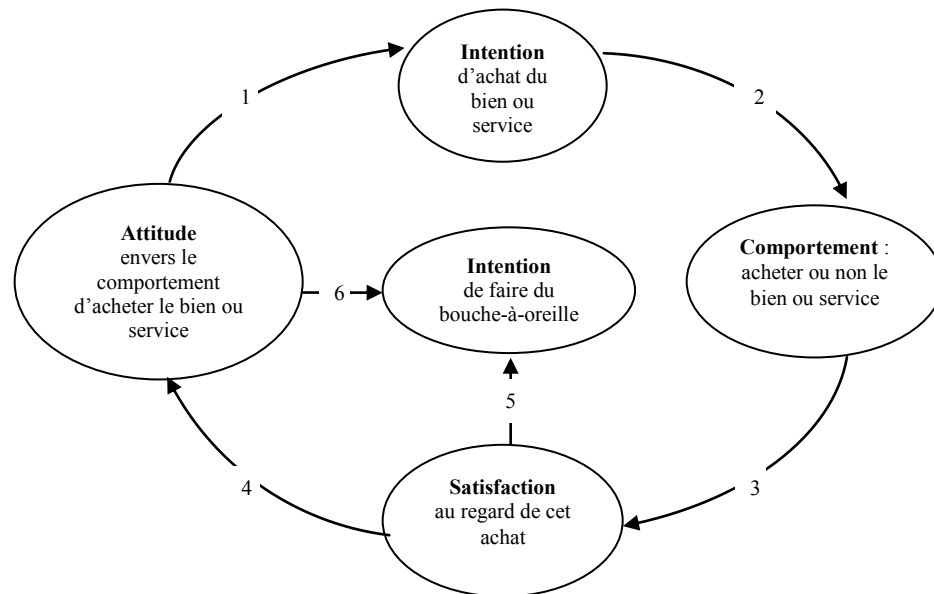


Figure 2.2 Le modèle théorique de la satisfaction des clients sur lequel repose l'étude

L'approche cognitive des méthodes de sondage

Avant les années 80, l'erreur dans un sondage s'estimait principalement selon un paradigme statistique (*i.e.*, l'erreur d'échantillonnage) et, dans une moindre mesure, selon un paradigme psychométrique (*i.e.*, l'erreur de mesure). Au cours des années 80, toutefois, un paradigme psychologique est apparu, complémentaire aux paradigmes statistique et psychométrique. Il est parfois connu sous l'acronyme CASM pour *Cognitive Aspects of Survey Methodology*, qui peut être traduit en français par les aspects cognitifs des méthodes de sondage. Ce nouveau courant d'études est né de l'application des concepts et des méthodes de la psychologie sociale et de la psychologie cognitive à l'étude des sources d'erreur dans un sondage. Le principal objectif de cette approche consiste à expliquer comment les gens décident s'ils acceptent de participer à un sondage, d'une part, et comment ils génèrent leurs réponses aux questions, d'autre part (Tourangeau, 2003).

Dans le cadre de la présente étude, c'est la façon selon laquelle les individus génèrent leurs réponses à un questionnaire qui est l'objet d'intérêt. Pour cette raison, le modèle du processus de réponse à un sondage de Tourangeau, Rips et Rasinski (2000) a été retenu dans notre étude. Tel que présenté au tableau 2.1, le modèle de Tourangeau et ses collaborateurs est composé de quatre étapes : la compréhension de la question, la récupération de l'information pertinente en mémoire, l'utilisation de cette information pour effectuer le jugement demandé et la réponse. Les auteurs précisent que l'ordre des étapes présenté ci-haut est probablement celui qui est le plus commun. Toutefois, ils mentionnent que les étapes peuvent se chevaucher, que l'ordre peut varier, que des retours à une étape précédente sont possibles, que certaines étapes peuvent être sautées, se dérouler en parallèle, ou encore être menées plus ou moins rondement. Voici plus précisément en quoi consiste le modèle de réponse à un sondage de Tourangeau et ses collaborateurs. Une plus grande importance sera accordée à la dernière composante du modèle, la réponse, puisqu'elle constitue l'étape d'intérêt dans la présente étude.

Tableau 2.1 Résumé des composantes du processus de réponse à une question de sondage (adapté de Tourangeau *et al.*, 2000, p. 8)

Composante	Processus spécifiques
Compréhension de la question	Prendre connaissance de la question et des instructions (s'il y a lieu).
	Identifier le sens de la question.
	Identifier quelle information on cherche à obtenir par cette question.
Récupération de l'information pertinente en mémoire	Récupérer l'information pertinente de la mémoire.
Jugement	Évaluer la précision des souvenirs éveillés.
	Au besoin, combler l'information partielle ou absente de la mémoire en l'inférant.
	Combiner en un tout cohérent l'information récupérée de la mémoire et celle inférée, et ce, de manière à former un jugement.
Réponse	Question fermée : identifier et sélectionner le choix de réponse qui correspond le plus au jugement formé.
	Question ouverte : rédiger la réponse (d'abord mentalement et ensuite à l'oral ou à l'écrit, selon le cas).

La compréhension de la question

Selon Tourangeau *et al.* (2000), la compréhension de la question constitue la première étape du processus de réponse à un sondage. Ces auteurs définissent la question comme étant une classe d'objets linguistiques typiquement utilisés pour demander de l'information (2000, pp. 26-27).

Cette première étape regroupe un ensemble de processus cognitifs. Dans un premier temps, le répondant doit porter attention à la question et, s'il y a lieu, aux instructions données. Certains problèmes peuvent dès lors survenir. Le répondant pourrait ne pas lire du tout les instructions ou ne les lire que partiellement. Dans un deuxième temps, le répondant identifie le sens de la question. Pour y arriver, il considère les mots utilisés, les unités linguistiques et la structure grammaticale. Ce sens littéraire est relativement stable entre les individus. Il demeure que certaines variables peuvent nuire à la compréhension du sens littéraire de la question. Par exemple, la syntaxe peut être ambiguë ou compliquée, les mots utilisés peuvent être vagues, non familiers, ou même trop techniques. Quant à la question en elle-même, elle peut avoir un double sens, inclure des présuppositions, ou contenir deux idées. Dans son interprétation de la question, le répondant peut également tenir compte du contexte dans lequel elle est posée, de ses connaissances sur le sujet et des choix de réponse offerts, le cas échéant. À titre d'exemple, la réponse à une question peut être influencée par les questions qui la précèdent (Tourangeau, Couper, & Conrad, 2004). En dernier lieu, le répondant ayant trouvé un sens à la question, il est alors en mesure de cerner quelle information est demandée.

La récupération de l'information pertinente en mémoire

Le répondant doit recourir à sa mémoire pour répondre à une question. La deuxième étape du processus de réponse à un sondage consiste donc en la récupération de l'information de la mémoire, information ayant bien sûr un lien significatif avec la question. Certains facteurs peuvent faire en sorte que le rappel de l'information en mémoire soit plus facile ou, au contraire, plus difficile (Jobe, Tourangeau, & Smith, 1993; Tourangeau *et al.*, 2000). Par exemple, le fait que la question demandée concerne un événement récent ou important dans la vie du répondant facilite le rappel de l'information cherchée. Il en est de même si la question demandée se rapporte à un événement atypique dans la vie du répondant, ou encore à quelque chose provoquant une forte charge émotionnelle chez ce dernier. Une fois que le répondant se rappelle les informations pertinentes à la question demandée, il passe normalement à l'étape du jugement.

Le jugement

La troisième étape du processus de réponse à un sondage porte sur le jugement. À cette étape, le répondant peut avoir un jugement clair en tête. C'est le cas lorsque le travail cognitif d'évaluation a déjà été fait et que, à l'étape précédente, le répondant a simplement retrouvé son jugement accessible en mémoire. Le répondant peut également avoir retrouvé en mémoire un jugement peu arrêté, partiel. Il peut aussi n'avoir jamais porté le jugement nécessaire pour répondre à la question et, en conséquence, n'avoir retrouvé aucun jugement pertinent en mémoire. Dans ces deux derniers cas, le répondant doit former son jugement en tout ou en partie avant de répondre à la question. Pour ce faire, il doit d'abord évaluer la précision des souvenirs pertinents éveillés à l'étape précédente. S'ils sont flous ou incomplets, il peut alors tenter de combler l'information manquante en l'inférant. Ensuite, les souvenirs éveillés et l'information inférée sont combinés en un tout cohérent. Enfin, le répondant forme son jugement à partir de ces informations. À ce sujet, Tourangeau et ses

collaborateurs mentionnent que plus le jugement du répondant est partiel ou inexistant dans sa mémoire, plus il risque d'être influencé par le contexte dans lequel la question est posée et d'être instable dans le temps.

En d'autres termes, dans le cas d'un sondage sur la satisfaction des clients, un répondant pourrait déjà avoir en mémoire une idée claire et précise de sa satisfaction envers un produit ou un service donné, mais il pourrait aussi n'avoir qu'une idée partielle ou vague de son degré de satisfaction envers ce produit ou ce service. Dans ce dernier cas, c'est à l'étape du jugement que le répondant va utiliser l'information qu'il a été capable de récupérer de sa mémoire pour déterminer son degré de satisfaction. Il arrive parfois que le répondant ait oublié certaines informations ou qu'il possède seulement une information partielle. Dans ce cas, il peut inférer les détails manquants, tel que déjà mentionné, ou combiner les fragments d'informations qu'il possède et construire son jugement sur cette base.

La réponse

La quatrième et dernière étape du processus de réponse à un sondage correspond à la réponse donnée. Les sondages utilisent deux grandes catégories de questions : les questions à réponse ouverte et les questions à réponse fermée. La première catégorie de questions laisse la liberté aux répondants de générer leurs propres réponses. Dans ce cas, la personne rédige mentalement sa réponse et la communique par la suite (Tourangeau *et al.*, 2000). Toutefois, ce sont généralement les questions à réponses fermées qui sont utilisées dans les sondages. Ce choix s'explique par le fait que les questions ouvertes, malgré la richesse d'information qu'elles peuvent apporter, coûtent plus cher à administrer, prennent plus de temps à analyser et sont plus difficiles à coder (Schwarz & Hippler, 1987). De plus, les choix de réponse peuvent amener le répondant à mieux saisir le sens de la question (Schwarz & Hippler, 1987; Strack & Martin, 1987) et à le guider quant aux types de réponses attendues (Hippler & Schwarz, 1987). Elles présentent cependant des

inconvenients. Pour répondre à une question à réponse fermée, la personne est contrainte d'utiliser des réponses déterminées à l'avance. De plus, les possibilités de réponses sont limitées aux choix offerts et ceux-ci qui ne correspondent pas nécessairement aux éléments de réponse auxquels la personne voudrait référer (Strack & Martin, 1987).

Les chercheurs utilisent souvent les questions à réponses fermées dans les études de satisfaction des clients. Plus précisément, le type de format de réponse qui nous intéresse est celui où les catégories de réponse sont ordonnées progressivement de manière à former une échelle de réponse. Il y a donc une gradation des choix de réponses de manière à ce que le premier choix représente, par exemple, le plus faible degré de satisfaction et que le dernier corresponde au degré le plus élevé de satisfaction.

Globalement, pour indiquer une opinion sur une échelle de réponse, les répondants tentent d'abord de situer leur jugement par rapport aux choix de réponse proposés. Ensuite, ils choisissent la réponse sur l'échelle qui correspond à leur jugement ou, du moins, celle qui s'en approche le plus (Tourangeau *et al.*, 2000). Plus précisément, la question à réponse fermée active deux types de représentations cognitives. Le premier type de représentations concerne le stimulus faisant l'objet de la question (p. ex. la courtoisie du personnel). L'autre type correspond aux représentations évoquées par l'échelle de réponse (p. ex. le degré de satisfaction). Chaque point de l'échelle active une structure cognitive particulière. Ces structures cognitives sont en quelque sorte des prototypes mentaux (Ostrom, 1987). Dans l'exemple mentionné, il s'agirait de prototypes mentaux de situations plus ou moins satisfaisantes et insatisfaisantes liés à chacun des points de l'échelle de réponse. La tâche du répondant est de trouver le meilleur appariement possible entre la représentation du stimulus associé à la question et les prototypes cognitifs activés par les choix de réponse (p. ex. à quel degré la courtoisie du personnel est considérée comme satisfaisante ou insatisfaisante). Parce que différents formats de réponse activent différentes structures cognitives, le format de réponse détermine en partie l'activité cognitive qui précède la réponse (Ostrom, 1987) et influence donc la réponse donnée. Il découle de ce raisonnement

théorique que les caractéristiques de l'échelle de réponse, telles que la numérotation des points, les étiquettes utilisées, la présentation graphique et le nombre de points utilisés peuvent toutes activer des représentations mentales différentes. En fonction de ces caractéristiques, chaque échelle de réponse peut donc être interprétée différemment par les répondants. En conséquence, les caractéristiques du format de réponse influencent l'endroit où les individus placent leur réponse sur l'échelle de mesure. Ceci a alors un impact direct sur la distribution ou la précision des réponses.

En somme, l'approche cognitive des méthodes de sondage postule que l'échelle de réponse n'est pas qu'un simple dispositif passif de mesure, elle serait également une source d'information pour le répondant (Schwarz, 1990, 1996; Schwarz & Hippler, 1987; Strack & Martin, 1987; Tourangeau, Couper, & Conrad, 2007; Tourangeau *et al.*, 2000). Les individus utilisent les caractéristiques de l'échelle de réponse pour assigner un sens aux différents points de l'échelle (Tourangeau *et al.*, 2007). De ce fait, elles n'activent pas forcément les mêmes structures cognitives et ne sont donc pas nécessairement interchangeables (Gannon & Ostrom, 1996). À titre d'exemples, la numérotation associée aux points de l'échelle (Amoo & Friedman, 2001; Schwarz, Grayson, & Knauper, 1998; Schwarz, Knauper, Hippler, Noelle-Neumann, & Clark, 1991), les étiquettes utilisées (Blais & Grondin, 2008; Wyatt & Meyers, 1987), l'orientation verticale ou horizontale de l'échelle (Stephenson & Herman, 2000) et certains éléments graphiques (Christian & Dillman, 2004; Smith, 1995; Tourangeau *et al.*, 2007) ont été identifiés comme des facteurs susceptibles d'influencer les réponses obtenues.

Dans le cadre de la présente étude, les deux caractéristiques d'intérêt sont la polarité et le nombre de points de l'échelle de réponse. En effet, rappelons-le, l'objectif général de cette étude est d'évaluer si la polarité et le nombre de points d'une échelle de réponse ont une influence sur les paramètres de distribution des réponses et sur les qualités psychométriques d'un questionnaire de satisfaction des clients et de ses items. Les prochains paragraphes présentent d'abord les études sur la relation entre la polarité d'une échelle de réponse et la

distribution des réponses. Ensuite, ils abordent les études ayant mis à l'épreuve le lien entre nombre de points dans l'échelle de réponse et les qualités psychométriques de l'outil de mesure.

Les études sur la relation entre la polarité de l'échelle de réponse et la distribution des réponses

Cette section de la thèse présente une analyse des études qui se sont intéressées à la relation entre la polarité d'une échelle de réponse et la distribution des réponses. Son objectif est double. Le premier est de présenter les constats qui découlent des études à propos de cette relation. Le deuxième est de vérifier si ces études appuient l'idée qu'une échelle de réponse unipolaire, contrairement à une échelle de réponse bipolaire, permettrait d'augmenter la variabilité des réponses de satisfaction rapportées par les clients envers un bien ou un produit donné et ainsi de corriger à la baisse leur degré général de satisfaction envers ce bien ou ce produit. Nous expliquons dans un premier temps les théories de Lam et Klockars (1982) et de Klockars et Yamagishi (1988) concernant l'effet de la polarité de l'échelle de réponse. Dans un second temps, nous présentons des études s'étant intéressées à l'effet de la polarité de l'échelle de réponse.

Lam et Klockars (1982) ont proposé une théorie pour expliquer par quel processus les individus font l'interprétation des différents points d'une échelle de réponse. Précisons que deux situations sont possibles. La première est celle où l'échelle de réponse n'est pas étiquetée, c'est-à-dire que seuls les deux points situés aux extrémités possèdent une étiquette. La deuxième est celle où l'échelle de réponse est étiquetée, c'est-à-dire que chacun de ses points possède une étiquette. Dans la situation où seules les deux extrémités de l'échelle de réponse possèdent une étiquette, les individus donneraient un sens aux points intermédiaires en divisant en parties approximativement égales la distance entre les deux points extrêmes. Dans le cas où une échelle de réponse possède une étiquette sur

chaque point, la signification donnée aux différents points varie selon les étiquettes utilisées. Si les points de l'échelle sont identifiés par des étiquettes créant des intervalles approximativement égaux, la distribution des réponses et la moyenne seront les mêmes que dans le cas d'une échelle où seules les extrémités possèdent une étiquette. Si les étiquettes sont formulées de telle sorte à créer des intervalles inégaux entre les points de l'échelle, la distribution des réponses différera par rapport à une échelle où seules les extrémités possèdent une étiquette. Par ailleurs, lorsque les étiquettes ne représentent que la partie positive du continuum psychologique considéré, le score moyen sera inférieur à celui d'une échelle de réponse où seuls les deux extrêmes sont identifiés. Lorsque les étiquettes ne représentent que la partie négative du continuum psychologique d'intérêt, le score moyen sera supérieur à celui d'une échelle de réponse où seuls les deux extrêmes sont identifiés.

Afin de mettre à l'épreuve leur théorie, Lam et Klockars (1982) ont préalablement déterminé, à l'aide d'un échantillon de 46 étudiants, les valeurs d'échelle² de 42 qualificatifs (mots ou expressions) pouvant potentiellement être utilisés comme étiquettes. Pour y arriver, ils se sont fortement inspirés de la méthode des intervalles d'apparence égaux de Thurstone et Chave (1929). Ensuite, ils ont développé quatre échelles de réponse en cinq points, chacune débutant par « *Poor* » et se terminant par « *Excellent* ». La première échelle possédait des étiquettes uniquement sur ses extrémités. Quant à la deuxième échelle, elle possédait des étiquettes sur chacun des points de l'échelle et celles-ci correspondaient à des intervalles approximativement égaux. Les troisième et quatrième échelles possédaient aussi des étiquettes sur chacun des points de réponse, mais celles-ci ne correspondaient pas à des intervalles égaux. Pour ces deux échelles, les étiquettes représentaient les parties positive et négative du continuum psychologique d'intérêt, respectivement. Il y avait donc quatre versions du questionnaire. Chacune était identique,

² Dans la méthode des intervalles d'apparence égaux de Thurstone et Chave (1929), la valeur d'échelle d'un qualificatif fait référence à sa position sur un continuum psychologique. Cette position est déterminée en demandant à un échantillon d'individus de situer le qualificatif sur le continuum psychologique à l'aide d'une échelle de onze points où 1 signifie que l'énoncé représente le degré le plus défavorable de ce continuum et 11 qu'il représente le degré le plus favorable du continuum. La valeur d'échelle du qualificatif correspond à la médiane des réponses données par les individus pour le situer sur le continuum psychologique.

seules les échelles de réponse variant. Les quatre versions du questionnaire ont été administrées aléatoirement à 375 étudiants. L'analyse de variance a montré qu'au moins une moyenne différait de manière statistiquement significative des autres. L'analyse *post-hoc* a permis d'établir que l'échelle dont les étiquettes ne représentaient que la partie négative du continuum psychologique présentait la moyenne la plus élevée ($M = 3,45$) et que celle dont les étiquettes ne représentaient que la partie positive du continuum psychologique avait la moyenne la plus faible ($M = 3,11$). Quant à la comparaison de l'échelle n'ayant pas d'étiquettes sur ses points intermédiaires avec celle ayant des étiquettes, les résultats indiquaient que les degrés de satisfaction rapportés n'étaient pas statistiquement différents, leurs moyennes étant toutes deux de 3,30. À la lumière de leurs résultats, les auteurs recommandent d'utiliser une échelle n'ayant des étiquettes qu'aux deux extrémités du continuum si l'objectif est de développer une échelle de mesure dont les intervalles sont apparemment égaux. Ils affirment aussi qu'il est sans doute possible d'obtenir une discrimination plus fine entre les sujets pour une partie du continuum psychologique en utilisant majoritairement des étiquettes représentant la partie du continuum pour laquelle une discrimination plus fine est souhaitée. Par exemple, pour augmenter le pouvoir de discrimination dans la partie positive du continuum psychologique, il suffirait d'utiliser majoritairement des étiquettes référant à cette partie du continuum.

Les résultats rapportés par Klockars et Yamagishi (1988) appuient ceux de Lam et Klockars (1982). Leur étude visait à évaluer si le sens donné par les répondants à différentes étiquettes varie en fonction de leur position³ sur l'échelle de réponse. Pour ce faire, ils se sont servis d'un questionnaire d'évaluation des professeurs à être administré à des étudiants. Les chercheurs ont développé cinq versions dudit questionnaire, chacune d'entre elles ne différait qu'en ce qui a trait à l'échelle de réponse (ces cinq échelles de réponse sont présentées à la figure 2.3). Pour chacune de ces échelles, les réponses minimale et maximale étaient « *1 = Very poor* » et « *6 = Excellent* ». Dans les trois

³ La position de l'étiquette réfère à son emplacement dans l'échelle de réponse. Par exemple, est-elle placée sur un point situé au début, au centre, ou à la fin de l'échelle de réponse ?

premières versions du questionnaire, les étiquettes « *Fair* » et « *Good* » se suivaient toujours dans l'échelle de réponse, mais n'étaient pas apposées sur les mêmes points; selon la version, elles étaient placées au début (version 1), au centre (version 2) ou à la fin de l'échelle (version 3). Dans la quatrième version du questionnaire, seules les extrémités de l'échelle de réponse possédaient une étiquette. Enfin, dans la cinquième version, l'échelle de réponse ne consistait pas en une échelle à proprement dit, mais simplement en une étiquette, par exemple « *Fair* » ou « *Excellent* », inscrite sur une ligne de réponse. Les cinq versions du questionnaire ont été distribuées à un échantillon de 164 étudiants de sorte que chacun d'eux reçoive l'une d'entre elles. Quant à la tâche des étudiants, elle n'était pas de répondre au questionnaire, celui-ci ayant été complété avant de leur être administré, mais d'évaluer le degré de favorabilité de chacun des choix de réponses inscrits dans le questionnaire, et ce, à l'aide d'une échelle allant de 1 « *Extremely unfavorable* » à « 15 = *Extremely favorable* ». À partir des évaluations des étudiants, les chercheurs ont calculé le score moyen de favorabilité pour chacun des points de chacune des cinq échelles de réponse. Après, ils ont comparé le degré de favorabilité des étiquettes « *Fair* » et « *Good* » selon leur position sur l'échelle de réponse. Les résultats obtenus suggèrent que l'interprétation d'un point de réponse est basée sur deux informations : l'étiquette qui y est associée et sa position dans l'échelle de réponse. À titre d'exemple, dans une échelle de réponse allant de « 1 = Très insatisfait » à « 5 = Très satisfait », l'étiquette « Satisfait » aurait une connotation plus péjorative pour les répondants si elle était située sur le deuxième point de réponse plutôt que sur le quatrième.

Échelle de réponse 1 : étiquettes « *Fair* » et « *Good* » placées au début

		<i>x</i>			
Very Poor	<i>Fair</i>	<i>Good</i>	Quite Good	Very Good	Excellent

Échelle de réponse 2 : étiquettes « *Fair* » et « *Good* » placées au centre

			<i>x</i>		
Very Poor	Poor	<i>Fair</i>	<i>Good</i>	Very Good	Excellent

Échelle de réponse 3 : étiquettes « *Fair* » et « *Good* » placées à la fin

				<i>x</i>	
Very Poor	Poor	So-So	<i>Fair</i>	<i>Good</i>	Excellent

Échelle de réponse 4 : étiquettes uniquement aux extrémités

Very Poor			<i>x</i>				Excellent
----------------------	--	--	----------	--	--	--	------------------

Échelle de réponse 5 : étiquette écrite sur une ligne de réponse

good

Figure 2.3 Les cinq échelles de réponse de l'étude de Klockars et Yamagishi (1988)

Il importe de souligner que quelques années plus tôt, l'étude exploratoire de Wildt et Mazis (1978), sans fondements théoriques particuliers, était arrivée à la même conclusion que celle Klockars et Yamagishi (1988). Ceux-ci cherchaient à déterminer, dans le contexte de l'évaluation d'un produit ou d'un service, si les individus considèrent uniquement les étiquettes sur l'échelle de réponse ou s'ils tiennent également compte de l'emplacement des étiquettes sur l'échelle de réponse. Leur démarche était similaire à celle employée par Lam et Klockars (1982). Tout d'abord, ils ont déterminé auprès d'un échantillon de 54 étudiants les valeurs d'échelle de 50 qualificatifs en s'inspirant de la méthode des intervalles d'apparence égaux de Thurstone et Chave (1929). À partir de ces qualificatifs, ils ont créé six échelles de réponse en cinq points dont les minimum et maximum étaient toujours

« *extremely poor* » et « *extremely good* », mais pour lesquelles les étiquettes sur les points intermédiaires différaient. Par la suite, ils ont élaboré six versions d'un questionnaire d'évaluation de produits et services, chacune intégrant l'une des six échelles de réponse. Enfin, ils ont administré aléatoirement les six versions du questionnaire à 479 étudiants. Leurs résultats montrent que la distribution des réponses varie selon les étiquettes utilisées et leur emplacement sur l'échelle de réponse.

Les conclusions des études de Lam et Klockars (1982), Klockars et Yamagishi (1988) ainsi que celles de Wildt et Mazis (1978) peuvent être appliquées au cas d'une mesure de satisfaction des clients. D'abord, puisqu'une échelle de réponse unipolaire utilise des étiquettes couvrant la partie positive du continuum de satisfaction, elle devrait davantage discriminer dans cette partie, c'est-à-dire là où la majorité des répondants se trouvent. Ensuite, une échelle de réponse unipolaire devrait résulter en une moyenne inférieure et une variance supérieure par rapport à une échelle de réponse bipolaire. Enfin, l'interprétation que fait un répondant d'un point d'une échelle de réponse dépend à la fois de l'étiquette qui y est attachée et de la position de ce point dans l'échelle de réponse, par exemple à son extrémité positive, au centre ou à son extrémité négative. Pour illustrer ceci, prenons l'exemple d'une échelle de réponse en cinq points. Supposons que l'étiquette « Assez satisfait » se retrouve au quatrième point de l'échelle dans la version bipolaire et au troisième point dans la version unipolaire. Selon la théorie de Lam et Klockars, l'étiquette « Assez satisfait » aura une connotation davantage négative dans la version unipolaire que dans la version bipolaire à cause de sa position relative dans l'échelle de réponse. Pour cette raison, la proportion d'individus endossant « Assez satisfait » ne sera pas la même selon qu'il s'agisse de l'échelle de réponse bipolaire ou unipolaire.

Maintenant que nous avons examiné les théories de Lam et Klockars (1982) et de Klockars et Yamagishi (1988) concernant l'effet de la polarité de l'échelle de réponse, nous présentons une analyse de sept études s'étant intéressées à cet effet, soit celles de Weiss, Davis, England et Lofquist (1967), Follman (1976), French-Lazovik et Gibson (1984),

O'Muircheartaigh, Gaskell et Wright (1995), Farmer, Thompson, Heil et Heil (2001), Davern et Cummins (2006) et Mazaheri et Theuns (2009). Cette analyse permettra d'avoir une meilleure idée de l'effet de la polarité de l'échelle de réponse sur les réponses obtenues à un questionnaire.

Les auteurs du *Minnesota Satisfaction Questionnaire* (MSQ) (Weiss, Davis, England, & Lofquist, 1967) ont proposé une méthode pour diminuer la dissymétrie négative des scores du MSQ, lequel évalue la satisfaction des individus envers leur emploi. Selon ce que rapporte le site Internet du laboratoire de recherche en psychologie industrielle de l'Université du Michigan (Vocational Psychology Research, 2007), pour atteindre cet objectif, ils ont remplacé l'échelle de réponse bipolaire par une échelle de réponse unipolaire. Les choix de réponse de départ (Très insatisfait, Insatisfait, Ni satisfait ni insatisfait, Satisfait et Très satisfait) ont respectivement été changées par ceux-ci : « Aucunement satisfait », « Très peu satisfait », « Satisfait », « Très satisfait », et « Extrêmement satisfait ». Le site Internet en question ne donne pas de précisions quant aux résultats obtenus, mais les chercheurs y mentionnent que ces modifications ont eu l'effet escompté, soit de diminuer la dissymétrie négative des scores. Une augmentation de la variance des items a également été rapportée. Il n'est toutefois pas spécifié si ces différences étaient statistiquement significatives et si elles ont résulté en une amélioration des qualités psychométriques de la mesure.

La recherche de Follman (1976) visait, quant à elle, à étudier deux caractéristiques de l'échelle de réponse, soit la catégorie d'étiquettes utilisée (variable indépendante à trois niveaux) et sa polarité (variable indépendante à deux niveaux), sur les réponses à un questionnaire d'évaluation des professeurs adressé à des étudiants ($n = 285$). Du croisement de ces deux facteurs, il est résulté six échelles de réponse à comparer entre elles. Chacune d'elles comptait cinq points de réponse et seules leurs extrémités possédaient une étiquette. Les trois catégories d'étiquettes comparées étaient celles reflétant (1) un degré d'accord ($1 = strongly disagree$ et $5 = strongly agree$), (2) une évaluation ($1 = inadequate$ et

5 = *excellent*) et (3) la nécessité d'une amélioration (*1 = needs great improvement* et *5 = needs no improvement*). Quant à la polarité de l'échelle de réponse, elle était soit bipolaire soit unipolaire. Les étiquettes aux extrémités des échelles de réponse bipolaires représentaient respectivement les parties négative et positive du continuum évalué (p. ex. *1 = strongly disagree* et *5 = strongly agree*), alors que celles pour les versions unipolaires représentaient uniquement la partie positive du continuum (p. ex. *1 = slightly agree* et *5 = strongly agree*). Les résultats de l'ANOVA visant à tester l'effet de ces deux facteurs sur la réponse moyenne au questionnaire indiquent que la catégorie d'étiquettes utilisée a une influence significative, mais pas la polarité. Plus précisément, les étiquettes reflétant la nécessité d'une amélioration (*1 = needs great improvement* et *5 = needs no improvement*) résultent en des moyennes supérieures aux deux autres types d'échelles. Cette étude ne supporte donc pas l'hypothèse que la polarité de l'échelle de réponse influence la distribution des réponses à un questionnaire.

L'étude de French-Lazovik et Gibson (1984) visait à éprouver l'hypothèse que la dissymétrie négative des scores observée dans les évaluations du personnel est attribuable à la bipolarité des échelles de réponse généralement employées. Plus précisément, ces auteurs soupçonnaient que les échelles de réponse bipolaires puissent faire en sorte que les items soient trop faciles à endosser pour les répondants. Ils ont donc testé leur hypothèse dans le cadre d'une évaluation des professeurs par leurs étudiants ($n = 304$). Le questionnaire d'évaluation comprenait quatre items. Dans la moitié des questionnaires administrés, les chercheurs ont utilisé une échelle de réponse bipolaire, dont les étiquettes étaient les suivantes : « *Very effective* », « *Above average* », « *Average* », « *Below Average* », « *Very Ineffective* ». Pour l'autre moitié des questionnaires, ils ont utilisé l'échelle de réponse unipolaire qui suit : « *Exceptional* », « *Very Good* », « *Good* », « *Satisfactory* », « *Unsatisfactory* ». À l'aide d'analyses de variance, les chercheurs ont vérifié si l'échelle de réponse employée affectait les moyennes des quatre items du questionnaire. Leurs résultats indiquent que les moyennes des réponses aux items sont inférieures de 0,273 à 0,458 pour les échelles de réponse unipolaires et que ces différences sont statistiquement significatives. Les auteurs mentionnent aussi que le degré d'asymétrie négative des items est moins

important si l'échelle de réponse est unipolaire, sans toutefois indiquer si cette différence est statistiquement significative. Par contre, les variances d'items ne diffèrent pas selon la polarité de l'échelle de réponse. Les auteurs concluent de ces résultats que l'échelle de réponse employée influence les moyennes d'items et leurs coefficients d'asymétrie. Ils ajoutent que l'utilisation d'une échelle de réponse unipolaire dans un questionnaire d'évaluation du personnel permet de diminuer la moyenne et l'asymétrie négative des réponses aux items et, conséquemment, de mieux distinguer entre elles les personnes évaluées.

L'un des objectifs de l'étude de O'Muircheartaigh, Gaskell et Wright (1995) était d'évaluer l'impact de la polarité des étiquettes associées aux extrémités d'une échelle de réponse à onze points sur la distribution des réponses. Seules les extrémités de l'échelle possédaient une étiquette. Les participants devaient répondre à la question suivante : « Est-ce qu'on devrait donner à l'*Advertising Standards Authority* plus de pouvoir pour contrôler la publicité ? », traduction libre de « *To what extent do you think the Advertising Standards Authority should be given more power to control advertisements ?* ». Une partie des participants ont répondu à cette question sur une échelle de réponse unipolaire allant de « *Not given any more power* » à « *Given much more power* », alors qu'une autre partie y a répondu sur une échelle bipolaire allant de « *Given much less power* » à « *Given much more power* ». L'expérimentation a eu lieu dans le cadre d'un sondage omnibus en Grande-Bretagne ($n = 2\ 165$). Les résultats de l'ANOVA indiquent que la polarité de l'échelle (unipolaire c. bipolaire) a une influence sur la distribution des réponses des sujets. En effet, la moyenne obtenue avec la version unipolaire est inférieure d'environ 0,5 point à celle obtenue avec la version bipolaire. En ce qui a trait à la variance, les tests C de Cochran et F de Bartlett-Box indiquent que les variances des deux versions d'échelles sont significativement différentes, c'est-à-dire plus grande d'environ 0,25 point (soit 0,5 en termes d'écart-type) pour la version présentant des étiquettes unipolaires. Les auteurs concluent de cette expérimentation que les étiquettes constituent une source d'information signifiante pour les répondants.

En ce qui a trait à l'étude de Farmer, Thompson, Heil et Heil (2001), elle consistait à évaluer si le fait de modifier les étiquettes apposées sur une échelle de réponse peut affecter la probabilité d'endosser un item. Une échelle de type Likert en cinq points était utilisée, tous les points possédant une étiquette. Le même questionnaire était passé à deux cohortes de sujets ($n_{\text{total}} = 8\,393$). Le questionnaire évaluait l'attitude des individus envers différents aspects de leur emploi. Les sujets composant ces deux cohortes répondaient au questionnaire à des moments différents, soit à deux années d'intervalle. Pour la deuxième cohorte, les étiquettes de l'échelle de réponse étaient modifiées pour 24 des 29 items du questionnaire. Pour 14 de ces 24 items, les changements concernaient la polarité de l'échelle de réponse ($1 = \textit{Not at all}$ à $5 = \textit{To a very great extent}$ comparativement à $1 = \textit{Strongly disagree}$ à $5 = \textit{Strongly agree}$), tandis que pour les 10 autres items, c'est l'étiquette correspondant au point milieu qui changeait (*neither disagree nor agree* comparativement à *neutral*). Pour vérifier si ces modifications avaient un effet sur la probabilité d'endosser les items, une analyse de fonctionnement différentiel d'items à l'aide du modèle logistique à deux paramètres de la théorie des réponses aux items fut menée. Chez les 10 items où le point milieu avait été changé, sept présentent un fonctionnement différentiel d'item significatif. Chez les 14 items où la polarité de l'échelle de réponse avait été modifiée, tous présentent un fonctionnement différentiel d'item significatif. Toutefois, contrairement à ce qu'avaient prévu les chercheurs, trois des cinq items pour lesquels l'échelle de réponse n'avait pas été changée affichaient un fonctionnement différentiel d'item significatif. Néanmoins, les fonctionnements différentiels d'items les plus importants ont été observés pour les items où la polarité de l'échelle de réponse avait été changée. C'est seulement pour ces derniers que des tailles d'effet (évaluées à l'aide de la statistique d de Cohen) supérieures ou égales à 0,30 ont été trouvées. Plus spécifiquement, les estimés du paramètre de difficulté pour les items dont l'échelle de réponse est unipolaire étaient supérieurs à ceux pour lesquels l'échelle de réponse est bipolaire. En d'autres mots, les items avec une échelle de réponse unipolaire étaient plus difficiles à endosser que ceux avec une échelle de réponse bipolaire, ce qui s'est traduit aussi par des moyennes significativement inférieures pour les items avec une échelle de réponse unipolaire plutôt que bipolaire. À titre d'exemple, en étant combiné à une échelle de réponse unipolaire, la moyenne et le paramètre de difficulté de l'item 21 se situaient respectivement à 2,42

($\dot{E}.-T. = 1,06$) et 0,849, alors qu'en étant combiné à une échelle de réponse bipolaire ces deux paramètres étaient de 2,82 ($\dot{E}.-T. = 1,11$) et de 0,380. Les auteurs ont conclu que le fait de modifier l'échelle de réponse a pu engendrer un fonctionnement différentiel d'item. Ils ajoutent que plus le changement apporté à l'échelle de réponse s'avère important, plus l'ampleur du fonctionnement différentiel d'item risque d'être grande. Les auteurs soulignent cependant que les résultats de leur étude peuvent être biaisés par des facteurs d'ordre méthodologique, notamment par l'utilisation de deux échantillons différents tirés à deux ans d'intervalle. En somme, cette étude apporte un appui partiel à l'hypothèse affirmant que la polarité de l'échelle de réponse peut influencer la distribution des réponses à un questionnaire. Selon ces résultats, l'influence de la polarité de l'échelle sur la distribution des réponses n'est peut-être pas majeure puisqu'un fonctionnement différentiel significatif d'item peut survenir sans que des modifications à l'échelle de réponse soient apportées.

L'étude de Davern et Cummins (2006), quant à elle, avait pour but d'explorer la relation entre la satisfaction et l'insatisfaction envers sa vie selon que ces jugements opposés sont mesurés à l'aide d'une échelle de réponse bipolaire ou de deux échelles de réponse unipolaires, l'une mesurant la satisfaction et l'autre mesurant l'insatisfaction. Pour ce faire, les auteurs ont développé deux versions d'un même questionnaire, soit les versions A et B. Dans la version A, les items mesurant la satisfaction et ceux mesurant l'insatisfaction avaient une échelle de réponse bipolaire qui allait de « *Completely dissatisfied* » à « *Completely satisfied* ». Dans la version B, les items mesurant la satisfaction avaient une échelle de réponse unipolaire portant uniquement sur les degrés de satisfaction, qui allait de « *Not at all satisfied* » à « *Completely satisfied* », alors que les items mesurant l'insatisfaction avaient une échelle de réponse unipolaire portant uniquement sur les degrés d'insatisfaction, qui allait de « *Not at all dissatisfied* » à « *Completely dissatisfied* ». Les échelles de réponse comptaient toutes onze points et seules leurs extrémités étaient étiquetées. La moitié des répondants a reçu la version A du questionnaire et l'autre moitié la version B ($n = 518$). Les résultats montrent que les réponses données aux items de satisfaction et d'insatisfaction à l'aide des échelles de réponses bipolaires sont corrélées de

manière positive, alors que celles données sur les échelles de réponse unipolaires sont corrélées de manière négative, c'est-à-dire que plus les scores de satisfaction sont faibles, plus les scores d'insatisfaction sont élevés et inversement. De surcroît, lorsque des échelles de réponse unipolaires sont utilisées, la somme des moyennes des scores de satisfaction et d'insatisfaction est approximativement de 100 %. Ces chercheurs concluent de ces résultats que la satisfaction et l'insatisfaction sont simplement la réciproque l'une de l'autre et que l'une ne peut augmenter sans que l'autre diminue. L'étude rapporte également que le degré moyen de satisfaction exprimé par les participants ne diffère pas que l'échelle de réponse utilisée soit bipolaire ou unipolaire (et portant uniquement sur la satisfaction).

Enfin, l'un des objectifs de l'étude de Mazaheri et Theuns (2009) consistait à évaluer la relation entre la polarité d'une échelle de réponse (bipolaire et unipolaire) et les réponses obtenues à deux items, le premier mesurant de manière globale le degré de satisfaction des gens envers leur vie (*i.e.*, mesure directe de satisfaction) et le deuxième mesurant de manière globale leur degré d'insatisfaction envers leur vie. Selon la condition expérimentale à laquelle les participants étaient assignés ($n = 1\,737$), ils devaient utiliser pour répondre à ces deux items une échelle de réponse bipolaire ou deux échelles de réponse unipolaires, c'est-à-dire une portant sur les degrés de satisfaction et l'autre portant sur les degrés d'insatisfaction. Les étiquettes sur les extrémités de l'échelle de réponse bipolaire étaient « *Very dissatisfied* » et « *Very satisfied* ». Les étiquettes sur l'échelle de réponse unipolaire mesurant la satisfaction étaient « *Not at all satisfied* » et « *Very satisfied* », et pour celle mesurant l'insatisfaction, ces étiquettes étaient « *Not at all dissatisfied* » et « *Very dissatisfied* ». Toutes les échelles de réponse comptaient onze points et seules leurs extrémités possédaient des étiquettes. Leurs résultats et leurs conclusions rejoignent ceux de Davern et Cummins (2006). En effet, les scores obtenus avec les échelles de réponse unipolaires d'insatisfaction et de satisfaction sont corrélés de manière importante ($-0,67 \leq r \leq -0,92$), c'est-à-dire que plus les scores sur l'échelle de réponse unipolaire de satisfaction sont faibles, plus les scores sur l'échelle de réponse unipolaire d'insatisfaction sont élevés et inversement. À partir de ces résultats, ces chercheurs concluent que la satisfaction et l'insatisfaction sont la réciproque l'une de l'autre et donc

que si l'une décroît l'autre croît automatiquement. De plus, les résultats d'une analyse de variance indiquent que le degré moyen de satisfaction exprimé par les participants ne varie pas selon la que l'on utilise une l'échelle de réponse bipolaire ou une échelle de réponse unipolaire mesurant uniquement la satisfaction. Les résultats de cette étude ne soutiennent donc pas l'idée qu'employer une échelle de réponse unipolaire ne mesurant que la partie positive du continuum de satisfaction, plutôt qu'une échelle de réponse bipolaire, influence la distribution des réponses à un questionnaire évaluant la satisfaction.

De toutes ces études se dégagent quelques constats. D'abord, à l'instar de Lam et Klockars (1982), quatre des sept études présentées ont noté que la polarité d'une échelle de réponse a un effet sur la distribution des réponses. Plus précisément, les moyennes des scores sont plus faibles lorsqu'une échelle de réponse unipolaire est utilisée plutôt qu'une bipolaire (*i.e.*, Farmer, Thompson, Heil, & Heil, 2001; French-Lazovik & Gibson, 1984; O'Muircheartaigh *et al.*, 1995; Vocational Psychology Research, 2007). Deux de ces études rapportent aussi des variances d'items plus élevées quand une échelle de réponse unipolaire plutôt qu'une bipolaire est utilisée (*i.e.*, O'Muircheartaigh *et al.*, 1995; Vocational Psychology Research, 2007).

Il ressort aussi que l'étiquetage des points de l'échelle de réponse pourrait modérer l'effet de la polarité. Plus spécifiquement, la polarité pourrait avoir un impact beaucoup moins important, voire nul, si seules les extrémités de l'échelle de réponse possèdent une étiquette. En effet, dans les trois études qui n'ont pas trouvé de lien entre la polarité de l'échelle de réponse et les réponses données par les participants (*i.e.*, Davern & Cummins, 2006; Follman, 1976; Mazaheri & Theuns, 2009), seules les deux extrémités de l'échelle de réponse possédaient une étiquette. En revanche, les points de l'échelle de réponse étaient tous étiquetés dans trois des quatre études (*i.e.*, Farmer *et al.*, 2001; French-Lazovik & Gibson, 1984; Vocational Psychology Research, 2007) où un tel lien a été identifié.

Finalement, bien que la littérature scientifique rapporte un effet de la polarité sur la moyenne et la variance des réponses des individus, l'ampleur de celui-ci est relativement faible. Par exemple, O'Muirheartaigh *et al.* (1995) rapportent des différences de moyennes et d'écart-types de l'ordre d'un demi-point d'échelle (sur une échelle de onze points). Les bénéfices au plan des qualités psychométriques d'utiliser une échelle de réponse unipolaire plutôt que bipolaire pour mesurer la satisfaction des clients demeurent par conséquent incertains.

Les études sur la relation entre le nombre de points de l'échelle de réponse et les qualités psychométriques d'un outil de mesure

Cette section examine les résultats des études qui ont évalué si l'utilisation d'une échelle de réponse de plus de sept points a un effet sur les qualités psychométriques d'un outil d'évaluation. Cet examen vise à vérifier si la littérature scientifique appuie l'idée qu'augmenter à environ dix le nombre de points de l'échelle de réponse d'un questionnaire de satisfaction des clients pourrait se traduire par augmentation de ses qualités psychométriques. Les études dans lesquelles la variable mesurée est la satisfaction sont d'abord présentées étant donnée leur pertinence par rapport à notre sujet de recherche. Les autres études sont discutées ensuite. Regrouper les études de cette manière permettra aussi de vérifier si les résultats diffèrent selon que la variable mesurée est la satisfaction ou une autre variable psychologique. Au préalable, il à noter qu'il y a longtemps que les chercheurs s'intéressent au nombre de points idéal à considérer dans une échelle de réponse. Un des premiers fut Likert (1932), qui proposa d'augmenter de deux à cinq le nombre de points sur l'échelle de réponse afin d'augmenter la fidélité d'une mesure psychologique. Depuis, le nombre de points mis à l'épreuve dans les études a varié de deux à 101, mais plus couramment de deux à onze.

Les études réalisées dans le contexte d'une mesure de satisfaction

Dans un premier temps, nous avons identifié trois études qui ont eu recours à une mesure du concept de satisfaction pour tester si l'utilisation d'une échelle de réponse de plus de sept points a un effet sur les qualités psychométriques d'un outil d'évaluation. Elles sont présentées ci-après par ordre chronologique.

Tout d'abord, en 1981, Neumann et Neumann ont étudié la relation entre le nombre de points dans l'échelle de réponse et la force des corrélations inter-items d'un questionnaire évaluant la satisfaction des étudiants envers les cours. Ils ont développé six versions de ce questionnaire qui différaient entre elles quant au nombre de points dans l'échelle de réponse, celui-ci pouvant être de deux, trois, quatre, cinq, sept ou dix points de réponse. Les questionnaires ont été administrés à un échantillon de 500 étudiants universitaires américains et chacun a reçu de façon aléatoire l'une ou l'autre des six versions du questionnaire. Les corrélations inter-items moyennes ont ensuite été calculées pour les six versions du questionnaire. Les moyennes respectives de ces corrélations sont de 0,14, 0,19, 0,37, 0,38, 0,47 et 0,47 pour les versions avec une échelle de réponse de deux, trois, quatre, cinq, sept et dix points. Les corrélations inter-items semblent donc augmenter avec le nombre de points dans l'échelle de réponse, mais uniquement jusqu'à sept points. En effet, la corrélation inter-items moyenne est la même peu importe que l'échelle de réponse compte sept ou dix points. La même conclusion peut être tirée pour la fidélité (coefficient alpha de Cronbach) du questionnaire de satisfaction, puisque cet indice psychométrique est directement lié à la force des corrélations entre les items d'un outil de mesure.

Plus tard, Alwin (1997) a réalisé une étude visant à comparer la fidélité et la validité d'un questionnaire de satisfaction envers sa qualité de vie selon que l'échelle de réponse comportait sept ou onze points. L'auteur a utilisé les données du *Quality of Life Survey* de 1978, une étude nationale menée à la grandeur des États-Unis ($n = 3\ 692$). L'intérêt

d'utiliser les données de cette étude était que les répondants devaient donner leurs réponses à la fois sur une échelle en sept points, allant de « *Satisfied* » à « *Dissatisfied* », et sur une échelle en onze points, présentée sous la forme d'un thermomètre. Ainsi, Alwin a pu comparer les réponses fournies à l'aide de ces deux échelles. Pour ce faire, l'auteur a réalisé deux analyses factorielles confirmatoires, l'une avec les items dont l'échelle de réponse compte sept points et l'autre avec les items dont l'échelle de réponse compte onze points. Les résultats montrent que 14 des 17 items sont plus fortement associés au trait latent lorsque l'échelle de réponse comporte onze points plutôt que sept. De même, pour la totalité des 17 items, les coefficients représentant l'influence de l'erreur de mesure aléatoire sont inférieurs lorsque l'échelle de réponse comporte onze points plutôt que sept. Cependant, aucune de ces différences n'a été éprouvée par un test statistique. L'auteur conclut quand même que son étude appuie la thèse selon laquelle le fait d'utiliser une échelle de réponse comportant un grand nombre de points, idéalement onze, favorise l'obtention d'outils de mesure avec de meilleurs estimés de fidélité et de validité.

Enfin, Coelho et Esteves (2007) ont évalué l'influence du nombre de points de l'échelle de réponse (cinq et dix points) sur la validité nomologique d'une mesure de la satisfaction des clients. Pour répondre à leur objectif de recherche, ils ont utilisé le *Customer Satisfaction Index*, un questionnaire administré par téléphone et mesurant sept variables : le degré de satisfaction envers différentes entreprises, le degré de fidélité envers ces différentes entreprises, l'image de ces entreprises, les attentes envers ces entreprises, la qualité des biens/services offerts par ces entreprises, la valeur (*i.e.*, le rapport qualité/prix) des biens/services offerts par ces entreprises et enfin la qualité de la gestion des plaintes faite par ces entreprises. L'enquête téléphonique a été réalisée au Portugal en 2004. La moitié des répondants ($n = 252$) a été assignée à la version du *Customer Satisfaction Index* dans lequel il leur était demandé d'exprimer leurs réponses sur une échelle à cinq points. L'autre moitié des répondants ($n = 253$) a été assignée à la version de ce même questionnaire où cette fois les échelles de réponse comportaient dix points. Les relations entre les sept variables du *Customer Satisfaction Index* ont été modélisées à l'aide d'un système d'équations structurelles. Ce modèle a été estimé deux fois, soit une fois avec les données

issues de la première version du questionnaire et une fois avec les données provenant de la deuxième version de ce même questionnaire. La validité nomologique a été évaluée à l'aide de deux indicateurs : le coefficient de détermination de la satisfaction des clients et le coefficient de détermination de la fidélité des clients (variable prédite par la satisfaction). Les résultats indiquent que ces deux indicateurs sont supérieurs dans le cas où des échelles de réponse en dix points avaient été utilisées. Cependant, ces différences n'ont pas été testées statistiquement. Plus précisément, le coefficient de détermination de la fidélité des clients est supérieur de 0,27 point (0,54 c. 0,27). Celui de la satisfaction des clients, par contre, n'est supérieur que de 0,07 point (0,66 c. 0,59). À partir de ces résultats, les auteurs concluent tout de même que l'utilisation d'échelles de réponse de dix points permet d'obtenir de meilleurs indices de validité nomologique, notamment dans le contexte d'un sondage portant sur la satisfaction des clients.

En résumé, sans démontrer hors de tout doute que des gains au plan des qualités psychométriques sont possibles en utilisant une échelle de réponse allant au-delà de sept points de réponse dans un questionnaire sur la satisfaction, les résultats de ces trois études permettent de le penser. En effet, les auteurs de deux des trois recherches examinées (*i.e.*, Alwin, 1997; Coelho & Esteves, 2007) arrivent à de meilleurs indices de validité et/ou de fidélité en augmentant l'échelle de réponse à dix (1 à 10) ou onze (0 à 10) points.

Les études non réalisées dans le contexte d'une mesure de satisfaction

Dans un deuxième temps, nous avons trouvé plusieurs études dans la littérature scientifique qui ont évalué si l'utilisation d'une échelle de réponse de plus de sept points peut avoir un effet sur les qualités psychométriques d'un outil d'évaluation. Nous synthétisons ici les résultats de ces études par rapport aux indices de fidélité et de validité d'un outil de mesure.

La grande majorité des études qui se sont intéressées à l'effet du nombre de points dans l'échelle de réponse sur la fidélité d'un outil de mesure ont abordé le concept de fidélité sous l'angle du degré de cohérence interne de la mesure tel qu'estimé par le coefficient alpha de Cronbach. À l'exception de l'étude de Mattel et Jacoby (1971), elles ont toutes trouvé une relation positive entre le nombre de points dans l'échelle de réponse et le coefficient alpha de Cronbach. Cependant, la relation entre le nombre de points dans l'échelle de réponse et l'importance du coefficient alpha de Cronbach n'apparaît pas linéaire, mais plutôt logarithmique (Alwin, 1992; Komorita & Graham, 1965; Lissitz & Green, 1975). En d'autres termes, chaque point de réponse supplémentaire engendre un gain en fidélité de moins en moins important. En ce sens, certains auteurs ont observé que le coefficient alpha de Cronbach augmente avec le nombre de points dans l'échelle de réponse jusqu'à cinq, six ou sept points, mais qu'il stagne au-delà de sept points (Bandalos & Enders, 1996; Birkett, 1986; Lissitz & Green, 1975; Preston & Colman, 2000). La même relation a aussi été observée entre le nombre de points dans l'échelle de réponse et le pouvoir de discrimination des items (corrélations item-total) (Preston & Colman, 2000). D'autres chercheurs rapportent que le degré de cohérence interne d'un outil de mesure s'accroît avec le nombre de points dans l'échelle de réponse jusqu'à neuf points (Alwin, 1992; Alwin & Krosnick, 1991; Weng, 2004), mais que cette augmentation est faible ou nulle à partir de dix points (Cox, 1980).

De surcroît, l'augmentation du coefficient alpha de Cronbach attribuable à l'utilisation d'un nombre supérieur de points dans l'échelle de réponse est relativement faible. À titre d'exemple, dans l'étude de Weng (2004), lorsque le nombre de points de réponse augmente de quatre à neuf, le coefficient alpha de Cronbach augmente de 0,72 à 0,79 dans la première expérimentation et de 0,79 à 0,82 dans de la seconde. De même, dans l'étude de Preston et Coleman (2000), le coefficient alpha de Cronbach passe de 0,81 à 0,86 lorsque le nombre de points dans l'échelle de réponse passe de deux à onze points. Rejoignant ces résultats, trois méta-analyses réalisées sur le sujet révèlent que le nombre de points dans l'échelle de réponse explique entre 1 % et 5 % de la variance du coefficient alpha de Cronbach (Bruner & Hensel, 1993; Churchill & Peter, 1984; Peterson, 1994).

Quant à la relation entre le nombre de points dans l'échelle de réponse et les indices de validité d'un outil de mesure, rares sont les études qui s'y sont intéressées, constat que Cox avait déjà fait en 1980. Soulignons que Mattel et Jacoby (1971) n'ont pas trouvé de différences aux plans des indices de validité concurrente et prédictive en fonction du nombre de points de réponse (qui allait de 2 à 19). Cependant, Preston et Coleman (2000) rapportent que la validité de convergence d'un outil de mesure de la qualité du service à la clientèle croît avec le nombre de points dans l'échelle de réponse jusqu'à six points, mais qu'elle demeure stable par la suite. En effet, dans cette recherche, l'indice de validité de convergence est respectivement de 0,83, 0,88 et de 0,88 lorsque l'échelle de réponse compte deux, six ou onze points de réponse.

En somme, dans les études sur la relation entre le nombre de points dans l'échelle de réponse et les qualités psychométriques d'un outil de mesure, les résultats diffèrent selon que celles-ci utilisent ou non une mesure de satisfaction. En effet, les études qui n'ont pas utilisé une mesure de la satisfaction concluent dans l'ensemble que peu ou pas de gains peuvent être réalisés en utilisant plus de sept points de réponse. En revanche, deux des trois études (*i.e.*, Alwin, 1997; Coelho & Esteves, 2007) qui ont utilisé une mesure de satisfaction ont trouvé que des gains pouvaient être faits en augmentant jusqu'à dix (1 à 10) ou onze (0 à 10) le nombre de points dans l'échelle de réponse. Ces dernières supportent donc la thèse qu'une mesure de satisfaction pourrait avoir de meilleures qualités psychométriques lorsque son échelle de réponse compte dix ou onze points par opposition à sept points ou moins.

Objectif spécifique et hypothèses de recherche

Afin de résumer les propos des deux premiers chapitres, rappelons qu'il est démontré depuis longtemps que la majorité des gens se disent satisfaits lorsqu'ils sont interrogés sur leur degré de satisfaction envers un bien ou un service (Danaher & Haddrell, 1996; Frank *et al.*, 1977; Larsen *et al.*, 1979; Linn, 1975; Nguyen *et al.*, 1983; Peterson & Wilson, 1992; Vanhamme, 2002; Westbrook & Oliver, 1981; Yi, 1990). Devant une telle problématique, certains auteurs ont suggéré d'utiliser une échelle de réponse unipolaire plutôt que bipolaire (French-Lazovik & Gibson, 1984; Gannon & Ostrom, 1996; Lam & Klockars, 1982; Weiss *et al.*, 1967) ou encore une échelle de réponse à dix (1 à 10) ou onze (0 à 10) points (Cummins, 2002, 2003; Cummins & Gullone, 2000; Fornell, 1992; Fornell *et al.*, 1996) afin d'augmenter la sensibilité des items de satisfaction et ainsi entraîner une amélioration des qualités psychométriques de l'échelle de satisfaction sous-jacente. En effet, selon ces auteurs, puisque les échelles de réponse unipolaires se concentrent uniquement sur la satisfaction et n'évaluent pas les degrés d'insatisfaction, chacun des items devrait être plus difficile à endosser, en plus de gagner en précision. De même, ils avancent qu'accroître à dix ou onze le nombre de points de l'échelle de réponse devrait permettre de mieux distinguer entre eux les individus quant à leur degré de satisfaction.

Décollant de ces propositions, l'objectif spécifique de cette étude est de mettre à l'épreuve l'effet de la polarité et du nombre de points de l'échelle de réponse sur la distribution des réponses aux items de satisfaction et sur les qualités psychométriques⁴ de l'échelle de satisfaction sous-jacente. À ces deux variables indépendantes nous en ajoutons une troisième à titre exploratoire, soit l'étiquetage des points de l'échelle de réponse, qui correspond à la présence ou l'absence d'étiquettes sur les points de l'échelle de réponse. En effet, tel que mentionné précédemment dans le cadre théorique, le fait d'identifier par des étiquettes chacun des points ou uniquement les deux extrémités de l'échelle de réponse

⁴ À l'instar de la grande majorité des études présentées dans le cadre théorique, cette thèse s'intéresse aux indices psychométriques classiques. Ce choix vise à faciliter la comparaison des résultats avec ces études antérieures.

pourrait avoir un effet modérateur (Baron & Kenny, 1986) sur l'effet de la polarité de l'échelle de réponse. Plus spécifiquement, l'effet de la polarité de l'échelle de réponse pourrait être moindre, voire nul, dans le cas des échelles de réponse où seules les extrémités ont une étiquette. Quelques études rapportent d'ailleurs que l'étiquetage des points de l'échelle de réponse est une variable qui peut affecter la moyenne des scores (Frisbie & Brandenburg, 1979), leur variance (Dixon, Bobo, & Stevick, 1984) ou la fidélité d'un instrument de mesure (Alwin & Krosnick, 1991; Weng, 2004). Par conséquent, la variable indépendante « étiquetage de l'échelle de réponse » a été ajoutée à cette recherche afin de tenir compte de cette possibilité. Comme l'effet de cette variable est étudié à titre exploratoire, aucune hypothèse formelle n'est émise à son sujet.

Nous émettons 15 hypothèses de recherche, les neuf premières concernent la polarité de l'échelle de réponse, alors que les six dernières ont trait au nombre de points de l'échelle de réponse. Plus précisément, par rapport à la polarité de l'échelle de réponse, nous émettons les trois hypothèses suivantes en lien avec la distribution des réponses aux items de satisfaction :

- H_1 : les items de satisfaction dont l'échelle de réponse est unipolaire présentent des indices de difficulté inférieurs à ceux des items dont l'échelle de réponse est bipolaire;
- H_2 : les items de satisfaction dont l'échelle de réponse est unipolaire présentent des écarts-types supérieurs aux items dont l'échelle de réponse est bipolaire;
- H_3 : les items de satisfaction dont l'échelle de réponse est unipolaire présentent des indices d'asymétrie moins importants que les items dont l'échelle de réponse est bipolaire.

En corollaire, nous émettons les deux hypothèses suivantes au regard du score global de satisfaction :

H_4 : la satisfaction exprimée dans un questionnaire dont l'échelle de réponse est unipolaire est moins élevée que celle exprimée dans un questionnaire dont l'échelle de réponse est bipolaire;

H_5 : la variance de la satisfaction exprimée dans un questionnaire dont l'échelle de réponse est unipolaire est plus grande que celle exprimée dans un questionnaire dont l'échelle de réponse est bipolaire.

Enfin, nous posons quatre hypothèses quant à l'association entre la polarité de l'échelle de réponse et les qualités psychométriques d'un questionnaire de satisfaction :

H_6 : les indices de discrimination des items de satisfaction dont l'échelle de réponse est unipolaire sont supérieurs à ceux dont l'échelle de réponse est bipolaire;

H_7 : le coefficient alpha de Cronbach d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celui d'un questionnaire dont l'échelle de réponse est bipolaire;

H_8 : le degré de validité de convergence d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celui d'un questionnaire dont l'échelle de réponse est bipolaire;

H_9 : le degré de validité nomologique d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celle d'un questionnaire dont l'échelle de réponse est bipolaire.

Avant d'aborder les hypothèses relatives au nombre de points de l'échelle de réponse, précisons que nous choisissons d'éprouver une échelle de réponse en dix points (allant de 1 à 10) plutôt qu'en onze points (allant de 0 à 10). Ce choix vise à faciliter la comparaison avec les échelles de réponse les plus courantes, soit celles comprenant cinq, six ou sept points, qui débutent généralement par 1 plutôt que par 0 (p. ex. de 1 à 5, de 1 à 6 ou de 1 à 7). En effet, les échelles de réponse en cinq, six ou sept points sont très populaires à cause des nombreuses recommandations en leur faveur (Cox, 1980; Likert, 1932; Neumann & Neumann, 1981; Tang, Shaw, & Vevea, 1999; Weng, 2004). Ce sont d'ailleurs les plus utilisées dans les études sur les comportements de consommation de produits et de services (Bruner & Hensel, 1993; Voss, Stem, & Fotopoulos, 2000). Comme base de comparaison à l'échelle de réponse en dix points, nous privilégions une échelle de réponse en six points (1 à 6). Cette décision vise à éliminer la source d'erreur qui pourrait résulter de la comparaison d'une échelle de réponse possédant un nombre pair de points (dix points) à une échelle de réponse en comprenant un nombre impair (cinq ou sept points) et, par conséquent, un point milieu (Si & Cullen, 1998).

Ainsi, à l'égard du nombre de points de l'échelle de réponse, nous émettons l'hypothèse suivante quant à la distribution des réponses aux items de satisfaction :

H_{10} : les items de satisfaction dont l'échelle de réponse a dix points présentent des coefficients de variation supérieurs aux items dont l'échelle de réponse a six points.

En corollaire, nous posons l'hypothèse qui suit relativement au score global de satisfaction :

H_{11} : la variance de la satisfaction exprimée dans un questionnaire dont l'échelle de réponses a dix points est plus élevée que celle obtenue dans un questionnaire dont l'échelle de réponse a six points.

Finalement, nous posons quatre hypothèses quant au lien entre le nombre de points dans l'échelle de réponse et les qualités psychométriques du questionnaire de satisfaction :

H_{12} : les indices de discrimination des items de satisfaction dont l'échelle de réponse a dix points sont supérieurs à ceux dont l'échelle de réponse a six points;

H_{13} : le coefficient alpha de Cronbach d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points.

H_{14} : le degré de validité de convergence d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points;

H_{15} : le degré de validité nomologique d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points.

En résumé, cette étude vise à évaluer l'effet de trois variables indépendantes, c'est-à-dire la polarité, le nombre de points et l'étiquetage de l'échelle de réponse, sur deux groupes de variables dépendantes, soit, premièrement, la distribution des réponses à des items de satisfaction et la distribution des scores globaux de satisfaction et, deuxièmement, les qualités psychométriques d'un questionnaire de satisfaction des clients. Quinze hypothèses seront vérifiées au total, permettant ainsi d'évaluer si l'emploi d'une échelle de réponse unipolaire plutôt que bipolaire ou en dix points plutôt qu'en six points peut augmenter la variance des scores de satisfaction et améliorer les qualités psychométriques d'un questionnaire de satisfaction. Nous étudierons aussi l'effet de l'étiquetage de l'échelle de réponse de manière exploratoire. De plus, contrairement aux recherches précédentes, la présente recherche éprouvera différentes combinaisons de ces trois facteurs. Au final, cette étude permettra d'avoir une meilleure compréhension de l'effet de ces variables dans le contexte d'un questionnaire de satisfaction et d'émettre à leur égard des recommandations pratiques ayant une valeur scientifique.

Chapitre 3 : Méthodologie

La présente partie de la thèse de doctorat présente la méthode employée afin de vérifier empiriquement les hypothèses de recherche énoncées à la fin du chapitre précédent. La stratégie de recherche et les décisions nécessaires à la réalisation de l'étude y sont explicitées. Les éléments abordés sont : le devis de recherche et ses variables, les participants, le questionnaire utilisé et, finalement, le plan d'analyse des données. Enfin, précisons que la collecte de données nécessaire à cette recherche s'est déroulée au Musée de la civilisation de Québec.

Variables et devis de recherche

Cette étude vise à mettre à l'épreuve l'effet de trois variables indépendantes, c'est-à-dire la polarité, le nombre de points et l'étiquetage de l'échelle de réponse, sur deux groupes de variables dépendantes, soit, premièrement, la distribution des réponses à des items de satisfaction et la distribution des scores globaux de satisfaction et, deuxièmement, les qualités psychométriques d'un questionnaire de satisfaction des clients. Plus précisément, nous nous intéressons aux relations entre ces variables indépendantes et dépendantes dans le cas d'un questionnaire multi-items de satisfaction. Effectivement, comme nous l'avons mentionné au premier chapitre, l'opérationnalisation de type multi-items du concept de satisfaction des clients est celle qui est la plus fréquente, et ce, tant la littérature scientifique que dans les entreprises.

Variabiles dépendantes

Le premier groupe de variables dépendantes est constitué d'indices statistiques décrivant la distribution de la satisfaction exprimée par les clients et concerne les hypothèses de recherche H_1 à H_5 ainsi que H_{10} et H_{11} . Dans un premier temps, pour évaluer l'effet des variables indépendantes sur la distribution des réponses aux items de satisfaction (H_1 , H_2 , H_3 et H_{10}), nous avons utilisé un indice de tendance centrale, soit leurs moyennes (H_1), des indices de dispersion, soit leurs écarts-types (H_2) et leurs coefficients de variation (H_{10}), de même qu'un indice d'asymétrie, c'est-à-dire leurs coefficients d'asymétrie (H_3). Précisons que pour vérifier H_2 , les écarts-types des items ont été préférés à leurs variances pour deux raisons. D'une part, parce l'écart-type est l'indice statistique généralement rapporté pour décrire la dispersion des scores autour de la moyenne. D'autre part, parce que l'écart-type est plus facile à interpréter que la variance étant donné qu'il est exprimé dans la même unité de mesure que la moyenne. Spécifions aussi que le coefficient de variation a été l'indice statistique privilégié pour comparer la variabilité des items de satisfaction en fonction du nombre de points de l'échelle de réponse (H_{10}) étant donné qu'il permet de comparer la variabilité de distributions qui n'ont pas la même unité de mesure. Dans un second temps, pour évaluer l'effet des variables indépendantes sur la distribution du score global de satisfaction, nous avons utilisé sa moyenne (H_4) et sa variance (H_5 et H_{11}). Dans ce cas-ci, la variance a été retenue étant donné qu'il était simple de comparer la variance du score global de satisfaction en fonction des différents niveaux des variables indépendantes à l'aide d'un test d'homogénéité des variances. À l'inverse, utiliser un test d'homogénéité des variances pour comparer la variance de chacun des multiples items du questionnaire de satisfaction en fonction des différents niveaux des variables indépendantes aurait nécessité une quantité trop importante de tests statistiques et augmenté de beaucoup le risque de commettre une erreur de type I. Pour cette raison, les indices de variabilité des items de satisfaction (*i.e.*, écarts-types, coefficients de variabilité) ont été comparés à l'aide de l'ANOVA, chaque item étant alors considéré comme une unité d'analyse, autrement dit un « sujet », dans la matrice de données. Le coefficient de variation et le coefficient d'asymétrie du score global de satisfaction n'ont pas été retenus comme variables dépendantes parce qu'il n'existe aucun test statistique, à notre connaissance, qui permet de

comparer entre eux deux ou plusieurs coefficients de variation ou coefficients d'asymétrie. En revanche, étant donné le grand nombre d'items de satisfaction, il était possible de comparer leurs coefficients de variation et leurs coefficients d'asymétrie en les considérant comme des unités d'analyse à l'intérieure d'une ANOVA.

Le deuxième groupe de variables dépendantes correspond aux qualités psychométriques du questionnaire de satisfaction des clients et réfère aux hypothèses de recherche H_6 à H_9 et H_{12} à H_{15} . Le pouvoir discriminant des items de satisfaction (H_6 et H_{12}) a été évalué à l'aide des indices de discrimination des items ou corrélations item-total corrigées. La fidélité du questionnaire de satisfaction (H_7 et H_{13}) a été estimée à l'aide du coefficient alpha de Cronbach. Quant à la validité conceptuelle du questionnaire de satisfaction des clients, elle a été évaluée à l'aide de deux stratégies. La première correspond au degré de validité de convergence (H_8 et H_{14}), estimé par la corrélation entre le questionnaire de satisfaction des visiteurs, opérationnalisé rappelons-le selon l'approche multi-items, et une mesure globale de satisfaction des clients. Quant à la deuxième stratégie, elle correspond au degré de validité nomologique (H_9 et H_{15}), mesuré par la force de l'association entre, d'une part, la satisfaction des clients et, d'autre part, les variables théoriquement liées selon le modèle présenté au deuxième chapitre, soit l'attitude envers le comportement de rachat et les intentions comportementales (*i.e.*, l'intention de rachat et l'intention de faire du bouche-à-oreille positif).

Variables indépendantes

Par ailleurs, trois variables indépendantes ont été étudiées. La première est la polarité de l'échelle de réponse, celle-ci pouvant être bipolaire ou unipolaire. Dans les échelles de réponse bipolaires, la moitié des choix de réponse expriment l'insatisfaction et l'autre moitié la satisfaction. Dans la présente étude, celles-ci allaient de « Extrêmement insatisfait » (1) à « Extrêmement satisfait » (6 ou 10, selon que l'échelle de réponse compte

six ou dix points). Pour leur part, les échelles unipolaires ne comprennent que des choix de réponse exprimant la satisfaction. Elles allaient de « Aucunement satisfait » (1) à « Extrêmement satisfait » (6 ou 10, selon que l'échelle de réponse compte six ou dix points). La deuxième variable indépendante est le nombre de points composant l'échelle de réponse, cette dernière comportant aussi deux niveaux : les échelles en six points et celles en dix points. La troisième variable indépendante à l'étude, l'étiquetage des points de l'échelle de réponse, comptait également deux niveaux. Ces deux niveaux sont la présence ou l'absence d'étiquettes sur les points de l'échelle de réponse, selon que tous les points ou uniquement les deux extrémités étaient identifiés par une étiquette, respectivement.

Le croisement de ces trois variables indépendantes à deux niveaux chacune, c'est-à-dire la polarité, le nombre de points et l'étiquetage de l'échelle de réponse, n'a pas résulté en huit échelles de réponse à éprouver, mais plutôt en six. La raison est qu'aucune échelle de réponse en dix points avec des étiquettes n'a été développée, car il est pratiquement impossible de générer des étiquettes pour chacun des points. Les six échelles de réponse qui ont été développées sont les suivantes : (1) bipolaire en six points (avec des étiquettes) (figure 3.1), (2) unipolaire en six points (avec des étiquettes) (figure 3.2), (3) bipolaire en six points (sans étiquettes) (figure 3.3), (4) unipolaire en six points (sans étiquettes) (figure 3.4), (5) bipolaire en dix points (figure 3.5) et (6) unipolaire en dix points (figure 3.6).

Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
1	2	3	4	5	6

Figure 3.1 L'échelle de réponse bipolaire en six points (avec des étiquettes)

Aucunement Satisfait	Très peu Satisfait	Passablement Satisfait	Assez Satisfait	Très Satisfait	Extrêmement Satisfait
1	2	3	4	5	6

Figure 3.2 L'échelle de réponse unipolaire en six points (avec des étiquettes)

Extrêmement Insatisfait						Extrêmement Satisfait
1	2	3	4	5		6

Figure 3.3 L'échelle de réponse bipolaire en six points (sans étiquettes)

Aucunement Satisfait						Extrêmement Satisfait
1	2	3	4	5		6

Figure 3.4 L'échelle de réponse unipolaire en six points (sans étiquettes)

Extrêmement Insatisfait										Extrêmement Satisfait
1	2	3	4	5	6	7	8	9	10	

Figure 3.5 L'échelle de réponse bipolaire en dix points

Aucunement Satisfait										Extrêmement Satisfait
1	2	3	4	5	6	7	8	9	10	

Figure 3.6 L'échelle de réponse unipolaire en dix points

Six versions du questionnaire ont donc été développées de manière à tester les six échelles de réponse à l'étude. Les items du questionnaire étaient les mêmes dans toutes les versions et seule l'échelle de réponse des items de satisfaction des clients variait d'une version à l'autre. La version 1 du questionnaire était constituée de l'échelle de réponse bipolaire en six points (avec des étiquettes) (figure 3.1), la version 2 de l'échelle unipolaire en six points (avec des étiquettes) (figure 3.2), la version 3 de l'échelle bipolaire en six points (sans étiquettes) (figure 3.3), la version 4 de l'échelle unipolaire en six points (sans étiquettes) (figure 3.4), la version 5 de l'échelle bipolaire en dix points (figure 3.5) et finalement la version 6 de l'échelle unipolaire en dix points (figure 3.6).

Devis de recherche

Un devis de recherche quasi-expérimental a été utilisé pour comparer les six échelles de réponse à éprouver. Il s'agissait plus précisément d'un plan factoriel incomplet à trois facteurs de deux niveaux chacun (*i.e.*, la polarité de l'échelle de réponse, son nombre de points et l'étiquetage de ses points). Ce devis résultait donc en six conditions différentes et chacune d'entre elles correspondait à une combinaison particulière des trois variables indépendantes étudiées. En clair, chaque condition correspondait à une échelle de réponse à éprouver.

Pour construire ce devis de recherche quasi-expérimental, six groupes de participants, ou groupes de comparaison, ont été créés. Pour y arriver, lors de la collecte de données au Musée de la civilisation, les six versions ont continuellement été distribuées en alternance aux participants, toujours de la version 1 à la version 6. De la sorte, à chaque tranche de six participants, chacune des versions du questionnaire avait été distribuée une fois, chacun des participants ayant répondu à l'une ou l'autre de ces six versions. De cette façon, à la fin d'une journée de collecte, un nombre approximativement égal de questionnaires avaient été complétés pour l'ensemble des six versions.

De plus, la procédure de collecte de données a été standardisée dans le but de favoriser son homogénéité. À cette fin, un court document (disponible à l'annexe 3) a été remis aux trois personnes qui ont travaillé sur le terrain pour solliciter les visiteurs du musée, c'est-à-dire le chercheur et deux assistants de recherche du Service de la recherche et de l'évaluation du Musée de la civilisation. Ce document contenait les directives à suivre pour réaliser la collecte de données, notamment sur la manière de solliciter les visiteurs. Lors d'une rencontre préparatoire avec les assistants de recherche, le chercheur s'est assuré que ces directives étaient comprises de la même façon par tous. Afin de vérifier si les assistants de recherche suivaient effectivement les directives données, le chercheur les a observés en

action lors de la première journée de la collection de données. Par la suite, un suivi téléphonique auprès des deux assistants de recherche a été fait au moins une fois par semaine, et ce, tout au long de la période de collecte de données.

Participants

Cette étude est principalement de nature méthodologique et psychométrique et ne vise pas à généraliser les résultats à une population très précise. Pour différentes raisons présentées ci-dessous, nous avons décidé d'administrer les questionnaires à des visiteurs adultes (18 ans et plus) du Musée de la civilisation, résidant au Québec et comprenant bien le français. Le choix de résidants exclusivement québécois est justifié par le fait que des études ont démontré que l'utilisation ou encore l'interprétation des échelles de réponse peut varier d'un pays à un autre (Bartikowski, Chandon, & Gierl, 2005; van Herk, Poortinga, & Verhallen, 2004). En conséquence, circonscrire la population-cible aux visiteurs Québécois permettait d'éliminer cette source de variance systématique. Ensuite, les moins de 18 ans ont été exclus parce que les études sur lesquelles la présente recherche prend appui ne concernent que des populations adultes. Utiliser des moins de 18 ans aurait introduit une source indésirable de variation et compliqué la comparaison de nos résultats avec ceux de la littérature. Enfin, la compréhension du français était nécessaire puisque seule une version en langue française du questionnaire avait été élaborée. Aucune version en langue anglaise du questionnaire n'a été développée, car même lorsque parfaitement bien traduits, les termes anglais peuvent avoir une signification quelque peu différente de leur équivalent français. En bref, la traduction du questionnaire aurait pu entraîner des fonctionnements différentiels d'items. Il aurait ensuite fallu effectuer des analyses séparées pour les questionnaires en anglais et ceux en français, réduisant ainsi la taille de l'échantillon et la puissance statistique des tests.

Une méthode d'échantillonnage de type non-probabiliste a été utilisée et la collecte des données s'est déroulée sur une période de près de deux mois, soit du 5 février au 2 avril 2009. La passation des questionnaires s'est faite lors de 27 des 50 journées où le musée était ouvert pendant cette période. C'est donc un peu plus d'une journée d'ouverture sur deux (*i.e.*, 54%) qui ont été couvertes par la collecte de données. Dans la mesure du possible, tous les visiteurs étaient sollicités à la fin de leur visite au Musée de la civilisation par l'une des trois personnes (deux hommes et une femme) assignées à cette fin. Approximativement 3 329 visiteurs du Musée de la civilisation ont été sollicités pour participer à l'étude, dont 1 762 correspondaient à la population-cible. De ce nombre, 1 008 ont accepté d'y participer, soit un taux de réponse de 57 %. Parmi ceux-ci, 101 ont été retirés de l'échantillon final étant donné qu'ils avaient omis de répondre à plus de 20 % des items de l'un ou l'autre des instruments de mesure composant le questionnaire.

L'échantillon final comptait 907 participants, soit approximativement 150 par groupe de comparaison (voir le tableau 3.1). Il était composé de 522 femmes (58 %) et de 366 hommes (40 %); 19 participants (2 %) n'ont pas indiqué leur sexe dans le questionnaire. Ces proportions sont similaires à celles obtenues dans l'Enquête générale de public de 2008, réalisée par le Musée de la civilisation, où 57 % des participants étaient de sexe féminin et 43 % de sexe masculin (Daignault, 2008). Les deux groupes d'âge les plus nombreux étaient les 25-34 ans (28 %) et les 35-44 ans (28 %), représentant ensemble plus de la moitié de l'échantillon. Le tableau 3.2 présente la répartition de l'échantillon par catégorie d'âge et par sexe.

Tableau 3.1 Nombre de participants dans chacun des six groupes de comparaison

Version du questionnaire	Nombre de questionnaires	
	complétés	%
1	141	15,6 %
2	152	16,8 %
3	154	17,0 %
4	153	16,9 %
5	157	17,3 %
6	150	16,5 %
Total	907	100 %

Tableau 3.2 Répartition par catégorie d'âge et par sexe des 907 sujets de l'échantillon*

Catégories d'âge	Sexe		Totaux
	Féminin	Masculin	
18-24 ans	79 (63,7 %)	45 (36,3 %)	124
25-34 ans	139 (56,5 %)	107 (43,5 %)	246
35-44 ans	151 (61,4 %)	95 (38,6 %)	246
45-54 ans	69 (55,2 %)	56 (44,8 %)	125
55-64 ans	61 (57,5 %)	45 (42,5 %)	106
65 ans et plus	22 (56,4 %)	17 (43,6 %)	39
Totaux	521 (58,8 %)	365 (41,2 %)	886*

*Pour 21 des 907 participants, le sexe et/ou l'âge sont inconnus.

Cette taille d'échantillon est apparue suffisante pour les besoins de l'étude. Elle permettait notamment d'avoir une bonne puissance statistique. À l'aide du logiciel GPower (Erdfelder, Faul, & Buchner, 1996), celle-ci a été estimée à 0,98 pour un test F omnibus comparant les six groupes au regard de leur score moyen de satisfaction (avec une taille d'effet estimée par ce même logiciel à 0,16). À titre de comparaison, la puissance

statistique minimalement acceptable est de 0,80 (Kirk, 1995). De plus, cette taille d'échantillon se compare avantageusement à celles des études portant sur l'effet de la polarité ou du nombre de points de l'échelle de réponse présentées dans le cadre théorique. La majorité des études empiriques relevées n'utilisait que quelques centaines de sujets, le plus souvent moins de 900. Par exemple, parmi les 17 études empiriques dont la taille d'échantillon était mentionnée, près des deux tiers (*i.e.*, 65 %) ont utilisé un échantillon de moins de 600 participants.

La participation des visiteurs à l'étude s'est faite sur une base volontaire et anonyme. La page couverture du questionnaire informait les participants du caractère confidentiel et anonyme des données recueillies (voir l'exemplaire du questionnaire à l'annexe 12). L'auteur de cette thèse, son directeur de thèse et le Musée de la civilisation sont les seuls à avoir accès aux données. Dans le but d'augmenter la motivation des visiteurs à y participer, ceux qui se portaient volontaires avaient droit à un bon de participation pour le tirage de quatre chèques-cadeaux d'une valeur de 50\$ chacun, valides dans les magasins Archambault. Le comité d'éthique de la recherche avec des êtres humains de l'Université Laval (CERUL) avait préalablement donné son accord à cette stratégie. Les bons de participation n'étaient pas rattachés aux questionnaires de sorte qu'il était impossible d'associer un bon de participation à un questionnaire. Ces bons ont tous été détruits une fois les prix remis aux quatre gagnants.

Questionnaire

Le questionnaire comprenait 71 questions présentées sur cinq pages en format légal (voir l'annexe 12)⁵. Les informations suivantes étaient données aux participants sur la première

⁵ La copie du questionnaire présentée à l'annexe 12 comprend huit pages étant donné qu'elle est formatée en format lettre plutôt qu'en format légal.

page : le titre de la recherche, la présentation du chercheur, le public ciblé par l'étude, le temps nécessaire pour répondre au questionnaire (approximativement 10 minutes), le caractère confidentiel et anonyme des réponses, le mode de consentement volontaire et implicite pour la participation à l'étude (par le retour du questionnaire) ainsi que le numéro d'approbation du projet de recherche par le Comité d'éthique de la recherche avec des êtres humains de l'Université Laval (CERUL).

Le contenu des quatre autres pages était divisé en quatre parties. Dans l'ordre, ces parties couvraient la satisfaction des visiteurs, leur attitude envers la fréquentation du Musée de la civilisation, leurs intentions comportementales par rapport au Musée de la civilisation et, enfin, leur profil comportemental muséal et sociodémographique. Plus précisément, la première partie comprenait les mesures multi-items et globale de satisfaction des visiteurs, tandis que la deuxième partie traitait de l'attitude envers la fréquentation du Musée de la civilisation. Quant à la troisième partie, elle comprenait deux mesures d'intention comportementale : l'intention de faire du bouche-à-oreille positif au regard de la fréquentation du Musée de la civilisation au cours du prochain mois et l'intention de revenir au Musée de la civilisation dans les six prochains mois. Enfin, la dernière partie dressait le profil du visiteur. Ce profil comprenait deux volets qui seront présentés dans la section suivante, soit le comportement de fréquentation du Musée de la civilisation et des musées en général et les variables sociodémographiques. L'élaboration des différents outils de mesure de même que leurs qualités psychométriques sont abordées par la suite.

Questions sur le profil des participants

Les questions sur le profil des participants se retrouvaient à la dernière page du questionnaire. Elles portaient sur leur comportement muséal, d'une part, et sur leur profil sociodémographique, d'autre part.

Comportement muséal

Trois questions mesuraient les comportements des répondants au regard de leur fréquentation du Musée de la civilisation et des musées en général. Les répondants devaient d'abord rapporter l'ensemble des expositions qu'ils avaient visitées dans la journée au Musée de la civilisation. Ensuite, ils devaient indiquer s'ils avaient déjà visité ce musée dans le passé. Dans l'affirmative, les répondants devaient indiquer : combien de fois ils étaient allés au Musée de la civilisation au cours de leur vie, le nombre de visites effectuées dans les six derniers mois et, enfin, le nombre de visites qu'ils font par année, en moyenne (choix de réponse : moins d'une fois par année, environ une fois par année ou plus d'une fois par année). La troisième question avait trait à leur fréquentation d'autres musées. Plus spécifiquement, les répondants devaient indiquer s'ils avaient déjà visité d'autres musées que le Musée de la civilisation et, dans l'affirmative, la fréquence moyenne annuelle à laquelle ils ont visité ces autres musées (choix de réponse : moins d'une fois par année, environ une fois par année ou plus d'une fois par année).

Profil sociodémographique

Des questions établissant le profil sociodémographique des répondants étaient présentées à la fin du questionnaire. Ces questions couvraient les variables suivantes : le sexe, l'âge, la provenance géographique, l'occupation principale et le niveau d'études. Elles étaient tirées de l'Enquête générale de public de l'été 2008 menée par le Musée de la civilisation (Daignault, 2008). Ces questions permettaient *a posteriori* de vérifier si chacun des participants faisait bien partie de la population-cible.

Les instruments de mesure et leur élaboration

Rappelons que selon le modèle théorique de la satisfaction des clients présentés au chapitre précédent, quatre variables devaient être mesurées : la satisfaction des clients, l'attitude envers le comportement de rachat, l'intention de rachat et l'intention de faire du bouche-à-oreille positif. Transposées au contexte du Musée de la civilisation, ces quatre variables devenaient : la satisfaction des visiteurs envers les expositions qu'ils ont vues dans la journée au Musée de la civilisation, leur attitude envers la fréquentation du Musée de la civilisation, leur intention de revenir au Musée de la civilisation et leur intention de faire du bouche-à-oreille positif à l'égard de la fréquentation du Musée de la civilisation.

L'élaboration des instruments mesurant ces variables s'est faite en suivant les règles de l'art en recherche psychométrique. La démarche s'est appuyée sur les étapes proposées dans la littérature scientifique en psychométrie (*e.g.*, Anastasi & Urbina, 1997; DeVellis, 1991; Fowler, 1995, 2002; Hambleton & Swaminathan, 1985; Nunnally, 1978). Les quatre grandes étapes ont consisté en (1) la définition opérationnelle des variables et la création d'une première version des instruments de mesure, (2) la vérification de la validité de contenu des instruments de mesure, (3) l'évaluation de la compréhension des items et de leur validité apparente par les individus de la population-cible et, finalement, (4) l'analyse empirique de la qualité de chacun des items des instruments de mesure.

Les prochains paragraphes décrivent ces quatre étapes. À l'intérieur de chacune de ces étapes, les quatre variables mesurées sont abordées à tour de rôle. Notons que l'élaboration des différents instruments de mesure y est décrite de façon détaillée. Il nous est apparu pertinent de procéder ainsi étant donné que le sujet de la thèse est de nature psychométrique. De plus, les items de satisfaction occupent une place centrale dans notre étude et, en conséquence, leur élaboration et leur sélection avaient une importance particulière.

Étape 1 : la définition opérationnelle des variables et la création d'une première version des instruments de mesure

Cette première étape visait à définir opérationnellement les différentes variables à mesurer, c'est-à-dire à identifier les indicateurs ou items qui allaient composer les instruments de mesure. Elle nécessitait aussi de déterminer le format des instruments de mesure (p. ex. les types d'items et d'échelles de réponse à utiliser). Elle s'est terminée par la génération de bassins d'items potentiels qui constituaient en fait les premières versions des instruments de mesure.

Satisfaction des visiteurs

Le concept de « satisfaction des visiteurs envers les expositions qu'ils ont vues dans la journée au Musée de la civilisation » a été opérationnalisé à la fois selon l'approche multi-items et selon l'approche globale. L'outil de mesure de type multi-items est celui qui nous intéresse. L'outil mesurant la satisfaction de manière globale vise, quant à lui, à estimer le degré de validité de convergence de l'outil de mesure multi-items. Afin d'alléger le texte, le concept de « satisfaction des visiteurs envers les expositions vues dans la journée au Musée de la civilisation » sera dorénavant appelé « satisfaction des visiteurs ».

Satisfaction des visiteurs : mesure multi-items

La mesure multi-items de la satisfaction des visiteurs nécessite que l'objet évalué, soit les expositions du Musée de la civilisation, soit décomposé en ses divers attributs. Les attributs sont l'ensemble des caractéristiques dudit objet pouvant être évaluées par le client et susceptibles d'influencer son degré de satisfaction. Pour construire cet outil de mesure, il fallait donc d'abord identifier les attributs que les individus évaluent lorsqu'ils visitent une

exposition au Musée de la civilisation. Les trois étapes que nous avons suivies pour y arriver sont l'examen de la littérature pertinente, notamment en évaluation muséale, puis la réalisation de deux études de nature qualitative, l'une avec une visée exploratoire et l'autre ayant une visée confirmatoire. Ces étapes sont décrites ci-après.

À l'étape première, une revue de la littérature portant sur l'évaluation muséale a été effectuée afin d'identifier les attributs d'une exposition muséale les plus significatifs et auxquels un visiteur réfère pour donner son degré de satisfaction. Dans le but d'identifier ces attributs les plus significatifs et par le fait même de favoriser la validité de contenu de l'échelle de satisfaction, seules les publications les plus récentes ont été retenues, c'est-à-dire celles réalisées entre 2000 et 2008. Toujours dans le souci de maximiser la validité de contenu, une étude antérieure à 2000 a tout de même été retenue. Il s'agit de l'étude réalisée en 1996 par le Groupe SECOR pour le Musée de la civilisation. Un des objectifs principaux de cette étude constituait à évaluer la satisfaction des visiteurs envers les expositions du Musée de la civilisation. Le fait qu'elle ait été réalisée au Musée de la civilisation plutôt que dans un autre musée la rendait d'autant plus pertinente. Dans la même veine, les attentes envers le Musée de la civilisation telles que rapportées par ses visiteurs dans l'Enquête générale de public de l'été 2008 (Daignault, 2008) ont aussi été considérées. Bien que ces dernières ne représentent pas précisément une mesure de la satisfaction des visiteurs, les attentes réfèrent néanmoins à des attributs jugés importants par le visiteur (Oliver, 1977, 1980; Spreng, MacKeinzie, & Olshavsky, 1996). Au total, 16 publications ont été retenues suite à cette première démarche. Celles-ci sont présentées au tableau A4.1 de l'annexe 4.

À la deuxième étape, une étude qualitative exploratoire a été menée auprès des visiteurs du Musée de la civilisation ayant les mêmes caractéristiques que notre population-cible, du moins sur la base des critères retenus pour circonscrire cette population. L'objectif de cette collecte de données consistait également à identifier les attributs d'une exposition muséale qui influencent le degré de satisfaction des visiteurs. Cette étude, en interrogeant

directement la population-cible, se voulait complémentaire à la revue de la littérature effectuée précédemment et permettait ainsi de favoriser la validité de contenu de la mesure multi-items de satisfaction.

Pour ce faire, un questionnaire (présenté à l'annexe 5) a été administré par le chercheur de l'étude aux participants à la sortie de l'une ou l'autre des expositions, selon le devis présenté ci-dessous (tableau 3.3). Il pouvait être auto-administré ou administré en face-à-face, selon la préférence du répondant. Ces deux modes d'administration du questionnaire ont été offerts suite aux recommandations de la chargée de recherche au service de l'évaluation et de la recherche du Musée de la civilisation. Selon son expérience des enquêtes muséales, dans le cas des questionnaires à questions ouvertes, le taux de participation est supérieur lorsqu'on offre aux visiteurs le choix entre ces deux modes d'administration du questionnaire, et ce, pour deux raisons. D'abord, certains visiteurs n'ont pas le goût de prendre le temps d'écrire des réponses. Ensuite, les visiteurs moins scolarisés préfèrent davantage exprimer leur réponse verbalement que par écrit.

Tableau 3.3 Devis de collecte des données de la deuxième étape de l'élaboration de la mesure multi-items de satisfaction des visiteurs (étude qualitative exploratoire)

Bloc horaire	Expos.*	Août			Septembre													
		Vend.	Sam.	Dim.	Lun.	Mar.	Merc.	Jeu.	Vend.	Sam.	Dim.	Lun.	Mar.	Merc.	Jeu.	Ven.	Sam.	
		29	30	31	1	2	3	4	5	6	7	8	9	10	11	12	13	
AM	1	x		x	x			x										
	2							x		x								
	3	x		x						x								
	4	x																
	5																	
	6																	
	7				x					x								
PM	1							x										
	2							x							x	x		
	3														x			
	4			x														
	5																x	
	6			x				x									x	
	7			x				x							x	x		

- * 1 Musée du quai Branly
 2 Or des Amériques
 3 Temps libre
 4 Urbanopolis
 5 Territoires
 6 Le temps des Québécois
 7 Nous, les premières nations

Plus spécifiquement, il était demandé aux participants d'indiquer oralement ou par écrit les éléments ou les facteurs qui font en sorte qu'ils ressortent satisfaits ou insatisfaits d'une exposition dans un musée. Les autres questions servaient à dresser le profil du visiteur, c'est-à-dire son comportement muséal et son profil sociodémographique. Cette étude s'est déroulée sur sept jours entre le 29 août et le 13 septembre 2008. Pour s'assurer d'avoir un échantillon de répondants le plus représentatif possible de la population-cible de notre étude, la collecte des données s'est faite le matin (de l'ouverture, soit 10h00, jusqu'à 13h00) et l'après-midi (de 13h00 à la fermeture, soit 17h00), les jours de semaine et de fin de semaine. En bref, la collecte des données couvre deux matins et deux après-midi de fin de semaine, ainsi que trois matins et deux après-midi de semaine.

Chaque avant-midi et après-midi, les visiteurs étaient sollicités pour participer à l'étude à leur sortie des salles d'exposition. En fait, pour tous les matins et pour tous les après-midi, de deux à quatre salles d'exposition avaient été identifiées. L'ordre dans lequel la tournée des expositions devait être faite était déterminé au hasard à l'aide d'un générateur de nombres aléatoires. Toutes les expositions ont ainsi été sélectionnées plusieurs fois. De même, pour éviter de ne solliciter systématiquement que les premiers visiteurs sortant d'une exposition, la stratégie suivante a été adoptée : un visiteur était sélectionné au hasard, à l'aide d'un générateur de nombres aléatoires, pour chaque tranche de 10 visiteurs. Des 124 visiteurs sollicités, 68 (55 %) correspondaient à notre population-cible. Parmi ceux-ci, 47 ont accepté de répondre au questionnaire (taux de réponse = 69 %). Les réponses des visiteurs, c'est-à-dire les attributs qu'ils ont rapporté évaluer lorsqu'ils visitent une exposition au Musée de la civilisation, sont présentées à l'annexe 6. Il est à noter que la saturation des données a été atteinte autour de 30 répondants. Un peu plus d'une dizaine de participants supplémentaires ont tout de même été ajoutés à l'échantillon pour être certain que la saturation des données avait bien été atteinte.

À la troisième étape, une seconde collecte de données à caractère qualitatif a été menée, mais cette fois avec visée confirmatoire plutôt qu'exploratoire. L'objectif de cette collecte de données consistait à s'assurer que la liste d'attributs identifiés lors des deux premières étapes était exhaustive. Cette liste a été élaborée en combinant les attributs identifiés à la suite de l'examen de la littérature et de l'étude qualitative exploratoire et contenait 74 attributs.

Un questionnaire auto-administré a été utilisé à cette troisième étape de la construction de l'échelle multi-items de satisfaction des visiteurs. À la première question, identique à celle du questionnaire utilisé à l'étape deux, il était demandé aux participants d'identifier les éléments qui font qu'ils ressentent satisfaits ou insatisfaits d'une exposition dans un musée. Cette question était à nouveau posée afin de vérifier si de nouveaux éléments propres à la période de l'année (à l'automne plutôt qu'à l'été) ne pouvaient pas être identifiés. Par la

suite, il était demandé aux participants de prendre connaissance de la liste des 74 attributs identifiés précédemment. Toutefois, afin de réduire la lourdeur de la tâche demandée aux participants, il a été décidé en cours de collecte de regrouper ces 74 attributs sous huit catégories et de leur présenter uniquement ces dernières accompagnées d'exemples. Le questionnaire présentant aux participants les huit catégories d'attributs plutôt qu'une liste exhaustive de ces derniers est présenté à l'annexe 7. Après avoir pris connaissance de ces attributs, les participants devaient indiquer si d'autres éléments que ceux présentés pourraient influencer leur degré de satisfaction envers une exposition du Musée de la civilisation. Dans la positive, ils devaient préciser ces attributs dans l'espace réservé à cette fin. La deuxième question, contrairement à la première, était donc plus susceptible de générer des attributs nouveaux puisque le fait de présenter des attributs, parfois accompagnés d'exemples, facilite le rappel d'éléments en mémoire par rapport à une question ouverte sans aucun stimulus pour faciliter le rappel (Churchland, 1992; Medin, Ross, & Markman, 2005).

La collecte de données de cette troisième étape s'est déroulée sur deux jours, soit une journée de fin de semaine (samedi le 22 novembre) et une journée de semaine (mercredi le 26 novembre), et ce, de l'ouverture à la fermeture du musée (10h00 à 17h00). Comme à la collecte de données précédente, l'ordre dans lequel la tournée des expositions devait être faite était déterminé au hasard à l'aide d'un générateur de nombres aléatoires. Toutes les expositions ont ainsi été sélectionnées plusieurs fois. De même, pour éviter de ne solliciter systématiquement que les premiers visiteurs qui sortaient d'une exposition, la stratégie suivante a été adoptée : un visiteur était sélectionné au hasard, à l'aide d'un générateur de nombres aléatoires, pour chaque tranche de cinq visiteurs. Des 53 visiteurs sollicités, 36 (68 %) correspondaient à la population-cible. Parmi ceux-ci, 21 ont accepté de répondre au questionnaire (taux de réponse = 58 %).

Les données collectées dans cette troisième étude ont permis de confirmer que la liste d'attributs identifiés suite à la revue de la littérature et à la collecte de données qualitative

exploratoire était exhaustive. L'identification des attributs que les individus évaluent lorsqu'ils visitent une exposition au Musée de la civilisation et qui influencent leur degré de satisfaction étant terminée, une première version de la mesure multi-items de la satisfaction des visiteurs a été élaborée. Cette version comprenait 72 items classés en six facteurs : (1) l'apprentissage et la réflexion (9 items), (2) le divertissement (10 items), (3) les thèmes des expositions (5 items), (4) le contenu des expositions (20 items), (5) les salles d'exposition (22 items) et (6) le personnel assigné aux salles d'exposition (6 items). Étant donné leur grand nombre, les items ont ainsi été regroupés pour que leur présentation soit visuellement moins lourde pour les répondants. Ces facteurs n'étaient pas le reflet d'une structure factorielle postulée *a priori*. Leur classement était basé sur la similitude de leur contenu et il a été obtenu en effectuant une analyse subjective des relations entre les items au regard de leur contenu. À ce stade de la recherche, nous ne pouvions faire autrement.

De manière plus concrète, les participants devaient répondre à la question « Considérant l'ensemble des expositions que vous avez vues aujourd'hui au Musée de la civilisation, quel est votre degré de satisfaction envers chacun des aspects suivants ? », en indiquant leur degré de satisfaction envers les aspects évalués par les 72 items. À titre d'exemple, il était demandé aux participants d'indiquer leur degré de satisfaction envers « le goût d'apprendre suscité par les expositions » (item 2), « le plaisir à visiter les expositions » (item 10), « la profondeur avec laquelle les thèmes des expositions sont traités » (item 22), « l'originalité du contenu des expositions » (item 27), « le matériel interactif » (item 41) et « l'esthétisme des salles d'exposition » (item 58). L'échelle de réponse allait de « 1 = Très insatisfait » à « 5 = Très satisfait ». Cette échelle de réponse conventionnelle a été utilisée puisque à cette étape de la recherche le choix des échelles de réponse à éprouver n'était pas encore arrêté. Le score de satisfaction de chacun des répondants était calculé en faisant la moyenne de leurs réponses aux 72 items. Ce score pouvait varier de 0 à 5. Les 72 items de cette version de la mesure multi-items de la satisfaction des visiteurs se retrouvent à l'annexe 8.

Satisfaction des visiteurs : mesure globale

Comme son nom l'indique, la mesure globale évalue la satisfaction envers l'objet d'évaluation dans son ensemble, sans égard direct à ses attributs. L'évaluation de la satisfaction envers l'objet considéré pris dans sa globalité ne nécessite qu'un seul ou très peu d'items (p. ex. deux ou trois). Tel que mentionné auparavant, ce type de mesure est souvent utilisé de manière complémentaire à la mesure multi-items. Il est aussi parfois utilisé seul dans les situations où le nombre de questions doit être limité. La mesure globale de satisfaction est aussi plus simple à élaborer étant donné qu'elle ne nécessite pas l'identification de l'ensemble des attributs de l'objet évalué pouvant affecter le degré de satisfaction des clients.

Il a été décidé de développer une mesure globale de satisfaction des visiteurs possédant plus d'un item dans l'objectif d'obtenir une fidélité supérieure. La mesure globale utilisée par les indices nationaux de satisfaction, comme l'*American Customer Satisfaction Index* (ACSI), a servi de base à l'élaboration de cette mesure. Ces indices permettent de mesurer la satisfaction des clients à l'aide de trois indicateurs ou items : (1) le client exprime, dans l'ensemble ou globalement, son degré de satisfaction envers le produit ou le service, (2) le client indique à quel point ses attentes ont été atteintes ou dépassées, et enfin (3) le client rapporte l'écart qu'il a perçu entre la qualité du service ou du produit qu'il a reçu et la situation idéale.

Les deux premiers indicateurs ont été retenus. Quant au troisième indicateur, il a été remplacé par un autre puisque l'offre idéale est un concept qui peut être difficile à concrétiser pour un répondant, notamment dans le cas où l'offre évaluée est complexe comme c'est le cas pour des expositions muséales. Ce troisième indicateur consiste en la valeur estimée de l'offre reçue, c'est-à-dire la perception du rapport qualité/prix. Cet indicateur nous est apparu un meilleur choix dans le présent contexte. D'abord, il est reconnu dans la littérature que cette variable est associée à la satisfaction des clients

(Athanasopoulos, 2000; Cronin, Brady, & Hult, 2000; Fornell *et al.*, 1996; Hallowell, 1996). Ensuite, le concept de rapport qualité/prix est bien connu dans la culture nord-américaine. La mesure globale de satisfaction peut être consultée à l'annexe 10. Elle correspond aux items 64 à 66. Ces trois items sont : (1) « Compte tenu de vos attentes de départ envers les expositions, quel est votre degré de satisfaction ? », (2) « Si vous considérez l'ensemble des coûts associés à une visite au Musée de la civilisation (ex. : déplacement, stationnement, billet d'entrée, etc.), quel est votre degré de satisfaction envers le rapport qualité/prix des expositions que vous avez visitées aujourd'hui ? » et (3) « Globalement, quel est votre degré de satisfaction envers les expositions que vous avez visitées aujourd'hui ? ». Tout comme pour la mesure multi-items, l'échelle de réponse allait de « 1 = Très insatisfait » à « 5 = Très satisfait ». De même, le score de satisfaction correspondait à la moyenne des réponses aux items et pouvait aller de 1 à 5.

Attitude envers la fréquentation du Musée de la civilisation

L'attitude envers la fréquentation du Musée de la civilisation a été opérationnalisée selon la méthode développée par Likert. Celle-ci consiste à élaborer des items qui reflètent clairement une attitude favorable ou défavorable envers le comportement d'intérêt et qui se répondent à l'aide d'une échelle de réponse exprimant divers degrés d'accord et de désaccord. L'échelle de réponse est codée de manière à ce que son minimum représente un fort degré de désaccord et son maximum un fort degré d'accord avec l'item. Ainsi, pour les items dont l'accord reflète une attitude favorable, un faible score à l'item indique une attitude défavorable et un score élevé, une attitude favorable. À l'inverse, pour les items dont l'accord reflète une attitude défavorable, un faible score à l'item indique une attitude favorable et un score élevé, une attitude défavorable. Pour ces derniers items, il est nécessaire de recoder les choix de réponse pour qu'ils puissent être interprétés sur le même continuum que les items reflétant une attitude positive. Par exemple, pour une échelle comprenant six points de réponse allant de 1 à 6, le 6 est recodé en 1, le 5 en 2 et ainsi de

suite. Le score d'attitude de chaque individu correspondant à la somme ou la moyenne des réponses aux items.

Pour élaborer cet outil de mesure selon la méthode proposée par Likert, les raisons pour fréquenter le Musée de la civilisation ont d'abord été identifiées. Pour ce faire, nous avons consulté les raisons qui sont ressorties de l'Enquête générale de public au Musée de la civilisation (Daignault, 2008). Selon ce rapport d'enquête, les personnes visiteraient le Musée de la civilisation pour se cultiver/découvrir de nouvelles choses, être surprises/attiser leur curiosité, être plongées dans l'ambiance d'un pays, d'une époque, se détendre/se divertir, passer un bon moment dans la convivialité, partager/échanger des idées, vivre une expérience hors du commun et participer à l'éducation des enfants. Parmi ces raisons, trois n'ont pas été retenues, soit « Être plongé dans l'ambiance d'un pays, d'une époque », « Partager/échanger des idées » et « Participer à l'éducation des enfants ». La première a été éliminée car ce ne sont pas toutes les expositions qui ont pour thème un pays ou une époque. La seconde a été supprimée car plusieurs visiteurs viennent seuls au Musée. Enfin, la troisième n'a pas été retenue étant donné que seulement une partie des visiteurs est accompagnée de leur ou leurs enfants.

Les six croyances conservées ont servi de point de départ pour élaborer une première liste d'items reflétant une attitude favorable ou défavorable envers la fréquentation du Musée de la civilisation. Au final, 26 items ont été élaborés, 19 reflétant une attitude favorable et sept reflétant une attitude défavorable envers la fréquentation du Musée de la civilisation. La liste complète de ces items se retrouve à l'annexe 9. À titre d'exemples, l'item « Pour moi, aller au Musée de la civilisation c'est relaxant » reflète une attitude favorable envers la fréquentation du Musée de la civilisation, alors que l'item « Pour moi, aller au Musée de la civilisation est une activité ennuyante » reflète une attitude défavorable. L'échelle de réponse possédait six points, numérotés de un à six, et dont les étiquettes étaient respectivement « Complètement en désaccord », « Assez en désaccord », « Légèrement en désaccord », « Légèrement d'accord », « Assez d'accord » et « Complètement d'accord ».

Le score d'attitude de chaque individu se calculait en faisant la moyenne des réponses aux items. Ce score pouvait donc varier entre 1 et 6.

Intention comportementale

L'intention comportementale est définie opérationnellement par Fishbein et Ajzen (1975) comme un type de croyances. Tout comme les croyances, l'intention est associée à un comportement donné. Elle constitue la probabilité subjective que l'individu adopte ledit comportement. En bref, l'intention comportementale correspond à la probabilité à laquelle un individu estime qu'il effectuera le comportement dont il est question. Cette probabilité subjective est mesurée à l'aide d'un ou de quelques items qui, par exemple, demandent aux répondants d'exprimer sur une échelle de réponse de type Likert s'ils ont l'intention ou non d'adopter ledit comportement (*e.g.*, Ajzen, 2002; Ajzen, Brown, & Carvajal, 2004; Ajzen, Brown, & Rosenthal, 1996; Fishbein & Ajzen, 1975; Fornell *et al.*, 1996; Hrubes, Ajzen, & Daigle, 2001; Oliver, 1980; Reinecke, Schmidt, & Ajzen, 1996).

Dans la présente étude, deux mesures d'intention ont été utilisées : l'intention de revenir au Musée de la civilisation et l'intention de faire du bouche-à-oreille positif incitant les gens à fréquenter le Musée de la civilisation.

Intention de revenir au Musée de la civilisation

L'intention de revenir au Musée de la civilisation est la première des deux mesures d'intention utilisée dans la présente étude. Nous avons choisi de mesurer l'intention des visiteurs de revenir au Musée de la civilisation lors des six prochains mois. Limiter la prévision de ce comportement à une période de temps relativement courte et rapprochée

avait pour objectif de faciliter la tâche des participants et la précision de leurs réponses. Spécifions qu'un intervalle de temps de six mois entre deux visites au Musée de la civilisation est relativement court. En effet, dans l'Enquête générale de public réalisée par le Musée de la civilisation (Daignault, 2008), seulement 38 % des participants ont rapporté que leur dernière visite au Musée avait eu lieu dans la dernière année.

Quatre items ont été élaborés pour mesurer ce concept : (1) « J'ai l'intention de revenir au Musée de la civilisation dans les six prochains mois », (2) « Je tiens fortement à revenir au Musée de la civilisation dans les six prochains mois », (3) « En pourcentage, quelle est la probabilité que vous reveniez au Musée de la civilisation dans les six prochains mois ? » et (4) « Environ combien de fois avez-vous l'intention de revenir au Musée de la civilisation dans les six prochains mois ? ». Le premier item mesure directement l'intention considérée et sert souvent d'indicateur d'intention comportementale (*e.g.*, Fornell *et al.*, 1996; Hrubes *et al.*, 2001; Reinecke *et al.*, 1996). Le deuxième item mesure le degré de motivation à revenir au Musée dans les six prochains mois. Les visiteurs devaient répondre à ces deux premiers items sur une échelle de type Likert en six points allant de « Complètement en désaccord » à « Complètement d'accord ». Le troisième item s'intéresse à la probabilité, exprimée en pourcentage, que le répondant revienne au Musée dans les six prochains mois. Les choix de réponse pour cet item étaient les suivants : 0-10 %, 11-20 %, 21-30 %, 31-40 %, 41-50 %, 51-60 %, 61-70 %, 71-80 %, 81-90 % et 91-100 %. Ceux-ci étant ensuite codés de un à dix. L'utilisation de cet item a déjà été rapportée dans la littérature (*e.g.*, Oliver, 1980) et il a été choisi étant donné qu'il reflète directement la définition de l'intention comportementale, qui est, rappelons-le, la probabilité subjective que l'individu effectue le comportement en question (Fishbein & Ajzen, 1975). Quant au quatrième item, il porte sur la fréquence prévue d'adoption du comportement. Ce dernier item était une question à réponse ouverte, c'est-à-dire que le participant devait rédiger sa propre réponse. Le score d'intention de revenir au Musée de la civilisation de chacun des participants correspondait à la somme des items. Nous aurions tout aussi bien pu prendre la moyenne, cela n'ayant pas d'influence sur l'estimation des paramètres du modèle d'équations structurelles (discuté plus loin), mais il aurait été plus difficile d'interpréter ce score

d'intention puisque deux des quatre items possédaient une échelle de réponse différente. Pour chaque répondant, ce score d'intention pouvait théoriquement osciller entre 3 et 205 (pour le cas extrême d'un individu qui dirait vouloir venir à tous les jours pendant les six prochains mois).

Intention de faire du bouche-à-oreille positif

L'intention de faire du bouche-à-oreille positif, rappelons-le, réfère à la communication interpersonnelle positive entre consommateurs à propos de l'objet de consommation. L'intention de faire du bouche-à-oreille positif (*i.e.*, la propension à recommander) qui pourrait inciter les gens à fréquenter le Musée de la civilisation a été opérationnalisée à l'aide des trois indicateurs utilisés par Zeithaml, Berry et Parasuraman (1996) et repris ultérieurement par d'autres auteurs (*e.g.*, Baker & Crompton, 2000; Cronin *et al.*, 2000; White & Yu, 2005), soit l'intention de recommander l'entreprise aux gens de son entourage, l'intention d'encourager les gens de son entourage à fréquenter l'entreprise et l'intention de faire des commentaires positifs aux gens de son entourage à propos de l'entreprise.

Le comportement a encore une fois été circonscrit dans le temps pour faciliter la tâche des participants et augmenter la précision de leurs réponses. C'est plus précisément l'intention de faire du bouche-à-oreille positif envers le Musée de la civilisation au cours du mois suivant la visite qui a été mesurée. Le choix d'un intervalle d'un mois repose sur l'hypothèse que c'est dans les jours qui suivent la visite au Musée que les personnes s'expriment le plus à leur entourage par rapport à leur expérience.

Ces trois items sont : (1) « Au cours du prochain mois, j'ai l'intention de recommander aux gens de mon entourage d'aller au Musée de la civilisation », (2) « Au cours du prochain

mois, j'ai l'intention d'encourager fortement les gens de mon entourage d'aller au Musée de la civilisation » et (3) « Au cours du prochain mois, j'ai l'intention de faire des commentaires positifs aux gens de mon entourage concernant le Musée de la civilisation ». L'échelle de réponse, de type Likert à six points, allait de « Complètement en désaccord » à « Complètement d'accord ». Le score des participants correspondait à la moyenne des trois items. Il pouvait ainsi varier entre 1 et 6.

Étape 2 : la vérification de la validité de contenu des instruments de mesure

La vérification de la validité de contenu des instruments de mesure visait à s'assurer que les items composant chacun d'eux couvraient bien les univers conceptuels visés. En ce sens, elle permettait d'identifier les items qui d'emblée n'apparaissaient pas être de bons indicateurs des variables mesurées, d'ajouter des items au besoin et d'améliorer leur formulation, si nécessaire. Cette évaluation a été faite par deux personnes, l'une experte de la mesure des attitudes, l'autre experte en évaluation muséale et employée au service de l'évaluation et de la recherche du Musée de la civilisation. La procédure employée avec ces deux experts fut de leur envoyer par courriel les instruments de mesure développés à l'étape précédente en leur demandant de les examiner et d'évaluer l'adéquation entre les items retenus et les concepts qu'ils sont censés mesurer. Aucune limite de temps ne leur était imposée. Les experts ont été rencontrés individuellement une fois leur travail d'évaluation terminé. Les modifications aux cinq outils de mesure suggérées par ces experts sont discutées ci-après. Précisons que l'ensemble de leurs recommandations a été suivi. Notons que les échelles de réponses, tout comme les valeurs possibles des scores totaux aux différents outils de mesure, n'ont pas été modifiés lors de cette deuxième étape.

Satisfaction des visiteurs : mesure multi-items

La première recommandation des experts concernant la mesure multi-items de la satisfaction des visiteurs était de réécrire les items dans une forme active, de sorte que ceux-ci soient plus simples à lire et à comprendre. Par exemple, après avoir reformulé les items, il était demandé aux répondants : (a) « Êtes-vous satisfait du degré auquel les expositions ont stimulé votre intérêt envers les thèmes traités », (b) « Êtes-vous satisfait de l'apport des expositions à votre culture générale ? », (c) « Êtes-vous satisfait de la quantité de choses nouvelles que vous avez découvertes ? » et ainsi de suite.

Les experts ont aussi recommandé de supprimer certains items et d'en combiner d'autres. L'objectif était de réduire quelque peu le nombre d'items en éliminant ceux qui d'emblée apparaissaient vagues, ambigus ou redondants. Plus précisément, l'item 9 (Les réflexions que suscitent chez vous les expositions) a été enlevé à cause de sa formulation lourde et du fait que l'item 8 (La qualité des réflexions proposées par les expositions) mesurait un contenu similaire. L'item 11 (La possibilité de passer un bon moment) a été supprimé étant donné qu'il est apparu ambigu aux yeux des juges et que deux autres items s'y apparentaient, soit les items 10 (Le plaisir à visiter les expositions) et 14 (Le côté divertissant de la visite des expositions). Parmi les cinq items suivants : « La capacité des expositions à vous impressionner, vous émerveiller » (item 13), « La capacité des expositions à vous faire vivre des émotions » (item 15), « La capacité des expositions à vous toucher personnellement » (item 16), « La possibilité de vivre une expérience hors du commun » (item 18), « La capacité des expositions à vous plonger dans une ambiance particulière » (item 19), seul l'item 15 a été conservé; il a été considéré le plus clair et résumant bien l'essentiel du contenu mesuré par les quatre autres items. L'item 23 (La capacité des expositions à toucher à plusieurs facettes des thèmes traités) a été retiré car il est apparu trop vague. L'item concernant « La facilité à comprendre les expositions » (item 26) a été supprimé étant donné qu'il était jugé ambigu. L'item 32, portant sur le degré de satisfaction envers « L'intérêt du contenu des expositions », a été supprimé étant donné son ambiguïté et le fait que d'autres items, tels que « L'originalité du contenu des expositions »

(item 27) ou « La possibilité de voir des choses que j'ai rarement ou jamais l'occasion de voir » (item 33), portaient également sur l'intérêt envers le contenu des expositions, tout en étant plus précis et plus clairs. Les items 50 (L'espace disponible pour contempler les objets exposés) et 54 (La possibilité de faire une visite agréable malgré l'achalandage) ont été retirés puisque d'autres items couvrant en bonne partie ce contenu leur ont été préférés, soit l'item 51 (L'espace disponible pour circuler dans les salles d'exposition) et l'item 52 (L'achalandage dans les salles d'expositions).

Les experts ont aussi suggéré de reformuler certains items pour les clarifier ou les simplifier. L'item 24 (La cohérence entre le thème et le contenu des expositions) a été modifié de la manière suivante : « Êtes-vous satisfait de la pertinence du contenu des expositions ? ». L'item 25 (La cohérence dans le contenu des expositions) a été retiré car il était redondant avec la version reformulée de l'item 24. Toujours dans l'objectif de clarifier et de simplifier certains des items, l'item 33 (La possibilité de voir des choses que j'ai rarement ou jamais l'occasion de voir) est devenu « Êtes-vous satisfait de la présence d'objets que vous avez rarement ou jamais l'occasion de voir ? » et l'item 34 (La possibilité de voir des objets historiques) a été remplacé par « Êtes-vous satisfait de la présence d'objets historiques ? ». Ensuite, le contenu de l'item 47 (L'éclairage dans les salles d'expositions) a été quelque peu clarifié en le reformulant de la façon suivante : « Êtes-vous satisfait de l'éclairage en général dans les expositions ? ». L'item 53 (Le respect du calme par les autres visiteurs) a été réécrit comme suit : « Êtes-vous satisfait du calme par les autres visiteurs dans les salles d'exposition ? ». La formulation des items 56, 66 et 67 a été simplifiée. L'item 56 (La quantité d'indications pour vous aider à vous diriger dans les expositions) a été changé pour « Êtes-vous satisfait des indications pour vous aider à vous diriger dans les expositions ? ». L'item 66 (La quantité de personnel disponible pour répondre à vos questions à l'intérieur des salles d'exposition) est devenu « Êtes-vous satisfait du nombre de personnes disponibles pour répondre à vos questions sur les expositions ? ». Enfin, l'item 67 (La capacité du personnel à l'intérieur des salles d'exposition à répondre à vos questions) a été remplacé par « Êtes-vous satisfait de la compétence du personnel pour répondre à vos questions sur les expositions ? ».

Certains items ont aussi été modifiés dans l'optique de respecter la terminologie muséale. Pour cette raison, d'une part, et pour être plus précis, d'autre part, les items 40 (Les possibilités d'interagir dans les expositions (jouer, toucher, voir, entendre, sentir, etc.)) et 41 (Le matériel interactif) ont été reformulés. L'item 40 est devenu « Êtes-vous satisfait du nombre d'interactifs ? » et l'item 41 « Êtes-vous satisfait de la qualité des interactifs (ex. : les jeux, le matériel informatique, etc.) ? ». L'item 39 (La quantité de contenu multimédia dans les expositions) a été retiré puisqu'il nous est apparu redondant avec les items 40 et 41. De même, l'expression « visites guidées », utilisée dans l'item 43, a été changée pour « visites commentées » et l'expression « fiches explicatives », utilisée dans les items 39, 45 et 46, a été remplacée par « vignettes d'objets », tout en prenant soin d'inscrire entre parenthèses la signification de ce terme, soit « les textes qui décrivent les objets ».

Certains items ont aussi été ajoutés suite aux recommandations des experts. Dans le but de mieux couvrir les types de contenus présents dans les expositions, l'item 42 (Les présentations audiovisuelles) a été scindé en deux items : « Êtes-vous satisfait du nombre de présentations audiovisuelles ? » et « Êtes-vous satisfait de la qualité des présentations audiovisuelles ? ». Aussi, les deux items suivants ont été créés pour évaluer la satisfaction des visiteurs envers les textes présentés dans les expositions : « Êtes-vous satisfait de la qualité des textes dans les expositions ? » et « Êtes-vous satisfait du nombre de textes dans les expositions ? », « Êtes-vous satisfait de l'éclairage des textes dans les expositions ? ». Enfin, les experts ont suggéré d'évaluer non seulement la satisfaction envers la quantité de sièges disponibles, mais aussi la satisfaction face à la qualité des sièges dans les salles d'exposition. L'item « Êtes-vous satisfait de la qualité des sièges mis à votre disposition pour vous reposer dans les salles d'exposition ? » a conséquemment été ajouté.

Une fois cette étape d'évaluation de contenu terminée, la mesure multi-items de la satisfaction des visiteurs comptait 63 items, toujours regroupés sous six facteurs. Le premier facteur (L'apprentissage et la réflexion) regroupait huit items, le second (Le divertissement) comptait cinq items, le troisième (Les thèmes des expositions) comprenait trois items, le quatrième (Le contenu des expositions) englobait 19 items, le cinquième (Les salles d'exposition) comportait 22 items et, finalement, le sixième facteur (Le personnel assigné aux salles d'exposition) chapeautait six items. Cette version peut être consultée à l'annexe 10 (questions 1 à 63).

Satisfaction des visiteurs : mesure globale

Les deux experts étaient d'accord pour dire que la mesure globale de satisfaction des visiteurs possédait une bonne validité de contenu. De plus, à leur avis, la formulation des items était appropriée. Cet outil de mesure fut donc conservé tel quel.

Attitude envers la fréquentation du Musée de la civilisation

Dans l'optique de réduire le nombre total de questions et par conséquent le temps de réponse au questionnaire, les experts ont suggéré d'éliminer certains items. Plus précisément, ils ont recommandé de supprimer les items 18 à 24 et de conserver l'item 13 (Aller au Musée de la civilisation me permet de voir des choses intéressantes), ce dernier couvrant le même contenu, mais de manière plus générale. L'item 14 (Aller au Musée de la civilisation me permet de découvrir des choses nouvelles pour moi) a été enlevé puisqu'il constitue un item factuel. Les items 15 (Aller au Musée de la civilisation est captivant) et 16 (Aller au Musée de la civilisation est stimulant) ont été enlevés car ils sont apparus trop vagues aux yeux des experts. Les items 14 (Aller au Musée de la civilisation me permet de découvrir des choses nouvelles pour moi) et 25 (Aller au Musée de la civilisation me

permet de voir des expositions dans un environnement convivial) ont également été jugés trop factuels par les experts. Les items 4 (Aller au Musée de la civilisation est une activité que je trouve trop intellectuelle) et 26 (Aller au Musée de la civilisation est une perte de temps) ont été retirés, les experts les jugeant trop extrêmes étant donné que le questionnaire s'adressait à des visiteurs du Musées de la civilisation, lesquels possèdent d'emblée une attitude favorable à la fréquentation de ce musée.

Au terme de cette démarche d'évaluation de la validité de contenu, 13 des 26 items élaborés à l'étape précédente, c'est-à-dire ceux présentés à l'annexe 9, ont été conservés. Les 13 items sont listés à l'annexe 10 (questions 67 à 79).

Intention de revenir au Musée de la civilisation

Les deux experts ont jugé que les quatre items mesurant l'intention de revenir au Musée de la civilisation au cours des six prochains mois étaient adéquats.

Intention de faire du bouche-à-oreille positif

Les experts ont considéré que les trois items de cet outil de mesure couvraient bien le concept d'intention de faire du bouche-à-oreille positif au regard de la fréquentation du Musée de la civilisation au cours du prochain mois.

Étape 3 : l'évaluation de la compréhension des items des instruments de mesure et de leur validité apparente par les individus de la population-cible

Avant de procéder à une analyse empirique des qualités psychométriques des instruments de mesure, une collecte de données a été réalisée auprès de la population-cible dans le but de s'assurer de la bonne compréhension des items et de leur validité apparente. La validité apparente réfère à la perception des répondants de la validité du test (Anastasi, 1988). C'est en quelque sorte une évaluation de la validité de contenu faite par des non-experts. Elle permet de vérifier si les répondants estiment que les items de l'instrument de mesure reflètent bien la variable que ces derniers sont censés mesurer. La validité apparente tire son importance du fait que si les répondants ne donnent pas de crédibilité à l'outil de mesure, leur motivation à bien répondre aux items sera affectée (Anastasi, 1988; Cronbach, 1969).

La méthode de collecte de données privilégiée a été de mener des entrevues cognitives (traduction libre de *cognitive interviews*). Cette méthode est suggérée par Fowler (1995) pour évaluer la compréhension des items par les individus de la population-cible. L'entrevue cognitive est une méthode qui consiste à sélectionner un petit échantillon de participants (maximum une dizaine) qui, dans le cadre d'entrevues individuelles, répondent aux différents items de l'instrument de mesure, d'une part, et expliquent à l'interviewer chacune de leurs réponses et ce qu'ils ont compris des questions, d'autre part. Les participants sont aussi invités à exprimer leurs incompréhensions et l'interviewer peut poser des questions aux participants pour clarifier leurs idées ou leurs commentaires (Willis, 2004). Comme l'explique Fowler (1995), cette technique permet de détecter si des items posent problème. Dans la positive, le chercheur peut immédiatement vérifier avec le répondant les nouvelles formulations qui pourraient éventuellement corriger cette ambiguïté.

L'évaluation de la validité apparente des cinq outils de mesure s'est faite lors des entrevues cognitives en demandant aux participants s'ils ont l'impression que la mesure évalue effectivement la variable qu'elle est censée mesurer. Par exemple, quand un participant avait terminé de répondre à la mesure multi-items de satisfaction des visiteurs, il lui était demandé s'il croyait que la mesure évaluait effectivement son degré de satisfaction envers les expositions visitées.

Ces entrevues ont eu lieu les vendredi et samedi 12 et 13 décembre 2009 entre 10h30 et 17h00. Suivant les recommandations de la responsable du service de l'évaluation et de la recherche du Musée de la civilisation, le chercheur a sollicité les visiteurs du musée une fois leur visite terminée. Il leur était demandé s'ils acceptaient de participer à l'évaluation d'un questionnaire portant sur la satisfaction des visiteurs, en ajoutant que cette étude faisait partie d'un projet de thèse de doctorat réalisé à l'Université Laval. Il était précisé que l'exercice prenait un maximum de 30 minutes et qu'en compensation pour leur temps il recevrait deux laissez-passer pour le Musée de la civilisation.

Lorsqu'une personne acceptait de participer à l'étude, le chercheur s'assurait d'abord qu'elle faisait partie de la population-cible. Il poursuivait ensuite en l'informant de l'objectif des entrevues cognitives, qui consiste à vérifier si les items du questionnaire sont bien compris par les visiteurs du musée. Par après, le chercheur expliquait au participant ce qu'il devait faire, soit (a) compléter le questionnaire (disponible à l'annexe 10), (b) exprimer sa compréhension des items, (c) expliquer ses réponses, (d) indiquer si des items lui posent problème et, finalement, (e) répondre aux questions qui lui seront posées par le chercheur au cours de l'entrevue.

Des 20 visiteurs sollicités, huit (cinq femmes et trois hommes) ont accepté de participer à l'étude, tous correspondant à la population-cible. Ce taux de réponse de 40 %, plus faible que ceux obtenus lors des deux études qualitatives réalisées à la première étape de

l'élaboration des instruments de mesure, s'explique par la durée des entrevues cognitives (près de 30 minutes) qui a semblé rebuter plusieurs des personnes sollicitées.

Les huit participants ont jugé que les outils de mesure évaluaient bien les concepts qu'ils étaient censés mesurer, autrement dit qu'ils possédaient une bonne validité apparente. Néanmoins, certaines remarques non en lien avec la validité apparente, mais concernant plutôt la « forme » que le « fond » de l'instrument de mesure, ont été faites par ces participants. Les paragraphes qui suivent font état de ces commentaires et des changements apportés aux instruments de mesure, le cas échéant. La mesure globale de satisfaction des visiteurs n'y est pas abordée puisqu'aucun participant n'a identifié de problèmes pour cet outil de mesure.

Satisfaction des visiteurs : mesure multi-items

En premier lieu, les huit participants ont mentionné que cette mesure comportait trop d'items. De leur point de vue, une vingtaine d'items, sur les 63, devraient être supprimés. Les commentaires des participants ont aidé à identifier puis supprimer les moins bons items, par exemple ceux redondants ou considérés non pertinents par les participants.

Les items « Êtes-vous satisfait de la quantité de choses que vous avez découvertes ? » (item 4), « Êtes-vous satisfait de ce que les expositions vous ont appris sur l'histoire ? » (item 5), « Êtes-vous satisfait de ce que les expositions vous ont appris sur les civilisations ? » (item 6) et « Êtes-vous satisfait de ce que les expositions vous ont appris sur le Québec ? » (item 7) ont été retirés. En effet, un participant a soulevé que le contenu évalué par ces items était déjà couvert de manière plus générale par l'item 1 (Êtes-vous satisfait de ce que vous avez appris dans les expositions) et l'item 3 (Êtes-vous satisfait de l'apport des expositions à votre culture générale ?). Ce participant a ajouté que les items 4 à

7 ne s'appliquaient pas à toutes les expositions, de sorte qu'il a trouvé difficile d'y répondre. Le même problème a été soulevé par ce participant pour l'item 22 (Êtes-vous satisfait de la présence d'objets historiques) qui a été supprimé. Un autre participant a soulevé que l'item 27 (Êtes-vous satisfait de la facilité à trouver rapidement l'information que vous cherchez dans les vignettes d'objet ?) était peu adapté au contexte du Musée de la civilisation étant donné que la plupart des vignettes d'objet sont relativement brèves. L'item a donc été retiré.

Les items « Êtes-vous satisfait de la fréquence des visites commentées ? » (item 34) et « Êtes-vous satisfait du nombre de fois où vous avez eu l'occasion d'utiliser un audio-guide (visite avec des écouteurs) ? » (item 35) n'ont pas été conservés. D'abord, sept des huit participants n'avaient pas souvenir des expositions qui offraient une visite commentée. Ensuite, les huit répondants ont indiqué ne pas avoir porté attention à la disponibilité d'audio-guides. De plus, il est fréquent que les expositions présentées n'offrent pas l'utilisation d'un audio-guide. L'item 59 (Êtes-vous satisfait de la compétence du personnel pour répondre à vos questions sur les expositions) et l'item 61 (Êtes-vous satisfait de la courtoisie du personnel dans les salles d'exposition) ont aussi été supprimés parce qu'ils ne s'appliquent pas à l'ensemble des participants. La moitié des huit participants ont dit ne pas être en mesure de donner une réponse à ces items étant donné qu'ils n'avaient pas interagi avec le personnel assigné aux salles d'exposition.

Dans un autre ordre d'idées, un participant a souligné que les items « Êtes-vous satisfait du degré auquel les expositions vous ont donné des émotions ? » (item 12) et « Êtes-vous satisfait de l'aspect captivant des expositions ? » (item 13) étaient difficiles à interpréter. L'item 57 (Êtes-vous satisfait de la qualité des sièges mis à votre disposition pour vous reposer dans les salles d'exposition ?) a posé problème à un autre participant. Ce dernier trouvait le terme « qualité » trop vague, d'une part, et l'item peu lié à son degré de satisfaction envers les expositions, d'autre part. Les items 12, 13 et 57 ont donc été retirés.

Les items « Êtes-vous satisfait de la facilité à repérer le parcours de visite dans les salles d'exposition ? » (item 46) et « Êtes-vous satisfait des indications pour vous aider à vous diriger dans les expositions ? » (item 47) étaient redondants, de l'avis d'un participant. À sa recommandation, seul l'item 47 a été conservé.

En deuxième lieu, les participants ont suggéré la reformulation de six items. L'item 19, « Êtes-vous satisfait de la présence d'objets qui vous ont émerveillé ? », a été reformulé après qu'un participant ait noté que cet item s'appliquait davantage à un musée d'arts qu'à un musée s'intéressant aux civilisations. Il a été remplacé par « Êtes-vous satisfait du nombre d'objets qui vous ont impressionné ? ». L'item 21 (Êtes-vous satisfait de la présence d'objets que vous avez rarement ou jamais l'occasion de voir ?), a été jugé ambigu par un répondant. Il a été changé pour « Êtes-vous satisfait du nombre d'objets que vous avez rarement ou jamais l'occasion de voir ailleurs ? ».

Un participant a trouvé que le terme « interactifs », lorsqu'employé sans explication (items 23 et 31), était difficile à comprendre, mais qu'il ne l'était pas en présence d'exemples, comme cela était le cas avec l'item 30 (Êtes-vous satisfait de la qualité des interactifs (ex. : les jeux, le matériel informatique, etc.) ?). Pour remédier à ce problème, l'item 23 (Êtes-vous satisfait de l'équilibre entre les différents types de contenus (ex. : textes, présentations audiovisuelles, objets, interactifs, etc.) ?) a été substitué par « Êtes-vous satisfait de l'équilibre entre les différents types de contenus (ex. : textes, présentations audiovisuelles, objets, jeux interactifs, etc.) ? » et l'item 31 (Êtes-vous satisfait du nombre d'interactifs) par « Êtes-vous satisfait du nombre d'interactifs (ex. : les jeux, le matériel informatique, etc.) ? ».

Un autre participant a mentionné que l'item 26 (Êtes-vous satisfait des explications fournies dans les vignettes d'objet (*i.e.*, les textes qui décrivent les objets) ?) serait plus clair et plus précis s'il était changé pour « Êtes-vous satisfait de la quantité d'information

dans les vignettes d'objet (*i.e.*, les textes qui décrivent les objets) ? ». Cette recommandation a été suivie. L'item 36 (Êtes-vous satisfait par rapport à la distance qui sépare les objets exposés de leur vignette d'objet (*i.e.*, les textes qui décrivent les objets) ?) était un peu ambigu selon l'avis d'un participant. Il a été changé pour « Êtes-vous satisfait de la facilité à trouver les textes descriptifs qui vont avec les objets ? ».

En somme, suite aux suggestions des participants aux entrevues cognitives, nous avons retiré 14 des 63 items de départ de la mesure de satisfaction des visiteurs de type multi-items. Nous avons aussi reformulé six des 49 items restants. Toutefois, pour réduire la partie du questionnaire portant sur la satisfaction des visiteurs à deux pages, tel que suggéré à l'unanimité par les répondants aux entrevues cognitives, quelques items supplémentaires devaient être retirés. Les deux experts ayant évalué la validité de contenu ont été consultés de nouveau pour cette tâche. Il leur a été demandé d'identifier les items dont le retrait nuirait le moins à la validité de contenu de l'outil de mesure.

Les experts ont suggéré de supprimer l'item 20 (Êtes-vous satisfait de la présence de belles choses à voir ?) puisque l'item 19 (Êtes-vous satisfait du nombre d'objets qui vous ont impressionné ?) couvre un contenu similaire et apparaît encore plus pertinent. Ils ont aussi proposé de retirer l'item 37 (Êtes-vous satisfait de l'éclairage des objets ?) puisque son contenu est couvert par l'item 42 (Êtes-vous satisfait de l'éclairage en général dans les salles d'expositions ?).

Toujours selon les recommandations des deux experts, les items « Êtes-vous satisfait de l'éclairage des vignettes d'objet ? » (item 39) et « Êtes-vous satisfait de l'éclairage des textes dans les expositions ? » (item 41) ont été combinés pour créer l'item suivant : « Êtes-vous satisfait de l'éclairage des textes en général (ex. : textes sur les murs, textes qui décrivent les objets, etc.) ? ». De même, les items « Êtes-vous satisfait de la lisibilité des vignettes d'objet ? » (item 3) et « Êtes-vous satisfait de la lisibilité des textes dans les

expositions ? » (item 40) ont été combinés pour produire l’item « Êtes-vous satisfait de la lisibilité des textes en général (ex. : textes sur les murs, textes qui décrivent les objets, etc.) ? ».

Enfin, en revisitant l’outil de mesure, un des experts a suggéré de réécrire l’item 45 (Êtes-vous satisfait du calme par les autres visiteurs dans les salles d’exposition ?) de la manière suivante : « Êtes-vous satisfait du calme dans les salles d’exposition ? ».

Cette troisième étape terminée, la mesure comptait alors 45 items, sept d’entre eux ayant été reformulés. Ces items correspondent aux questions 1 à 45 du questionnaire⁶ qui a été utilisé pour la mise à l’épreuve empirique des qualités psychométriques des cinq outils de mesures développés, soit la quatrième étape du processus d’élaboration des instruments de mesure. Celui-ci peut être consulté à l’annexe 11.

Attitude envers la fréquentation du Musée de la civilisation

Un seul item de la mesure de l’attitude envers la fréquentation du Musée de la civilisation a suscité des remarques. Il s’agit de l’item « Pour moi, aller au Musée de la civilisation c’est compliqué » (item 73). Trois participants ont mentionné qu’ils ne comprenaient pas cet item et que celui-ci devait être précisé. À partir de leurs suggestions, cet item a été reformulé de la manière suivante : « Pour moi, aller au Musée de la civilisation c’est

⁶ Dans ce questionnaire, l’échelle de réponse utilisée pour les deux mesures de satisfaction des visiteurs comportait six points plutôt que cinq, identifiés par les étiquettes suivantes : (1) Très insatisfait, (2) Assez insatisfait, (3) Légèrement insatisfait, (4) Légèrement satisfait, (5) Assez satisfait et (6) Très satisfait. En effet, c’est à ce stade du projet de recherche qu’il a été décidé qu’une échelle de réponse en six points servirait de base de comparaison à l’échelle de réponse en dix points. Suite à ce changement, les scores possibles aux mesures multi-items et globale de satisfaction pouvaient désormais prendre une valeur entre 1,00 et 6,00.

compliqué (*i.e.*, accessibilité, heures d'ouverture, etc.) ». Cet outil de mesure n'a pas subi d'autres changements à cette étape.

Intention de revenir au Musée de la civilisation

Les entrevues cognitives ont révélé que les items de la mesure de l'intention de revenir au Musée de la civilisation étaient clairs. Cependant, l'item 84 (Je tiens fortement à revenir au Musée de la civilisation dans les six prochains mois) a été jugé trop extrême par trois participants. Il a néanmoins été conservé pour le prétest à la suggestion des experts.

Intention de faire du bouche-à-oreille positif

Comme pour la mesure précédente, les entrevues cognitives ont révélé que les items de l'outil de mesure de l'intention de faire du bouche-à-oreille positif étaient bien compris des participants. Cinq d'entre eux ont jugé trop extrême et quelque peu irréaliste l'item « Au cours du prochain mois, j'ai l'intention d'encourager fortement les gens de mon entourage d'aller au Musée de la civilisation » (item 81). Étant donné que plus de la moitié des répondants ont fait cette remarque, cet item n'a pas été conservé pour le prétest, portant cette mesure à deux items.

Étape 4 : l'analyse empirique de la qualité des items des instruments de mesure

L'analyse empirique de la qualité des items de chacun des instruments de mesure du questionnaire s'est faite dans un contexte qui était le plus près possible de celui de l'étude pour laquelle ils étaient développés. Cette collecte de données visait à apprécier de manière générale les qualités psychométriques des différents instruments de mesure. Le questionnaire utilisé lors de ce prétest est disponible à l'annexe 11. Les cinq outils de mesure y apparaissent dans l'ordre suivant : les mesures multi-items (questions 1 à 45) et globale (questions 46 à 48) de satisfaction des visiteurs, la mesure d'attitude envers la fréquentation du Musée de la civilisation (questions 49 à 61), les mesures d'intention de faire du bouche-à-oreille positif au regard de la fréquentation du Musée de la civilisation (questions 62 et 63) et d'intention de revenir au Musée de la civilisation (questions 64 à 67).

Pour réaliser cette collecte de données, des visiteurs du Musée de la civilisation ont été recrutés pendant cinq journées complètes entre le 23 décembre 2008 et le 1^{er} janvier 2009. La méthode d'échantillonnage retenue est la même que celle utilisée dans l'étude principale. Les visiteurs qui se dirigeaient vers la sortie ou vers le vestiaire pour récupérer leur manteau étaient sollicités. Des 149 visiteurs sollicités pour répondre au questionnaire, 103 répondaient aux critères établis pour faire partie de la population-cible (69 %). Sur ces 103 visiteurs, 69 ont accepté de répondre au questionnaire (taux de réponse = 67 %). Quatre questionnaires ont dû être éliminés puisqu'ils n'avaient été que partiellement répondus. L'échantillon final sur lequel les analyses psychométriques préliminaires ont pu être réalisées comptait donc 65 participants.

L'examen du tableau 3.4 révèle que les instruments de mesure semblaient posséder une bonne validité nomologique. En effet, les construits sont tous corrélés entre eux de manière statistiquement significative (les corrélations variant entre 0,38 et 0,52), tel que suggéré par

le modèle théorique retenu pour notre étude et décrit précédemment. De plus, la mesure multi-items de satisfaction corrélait avec la mesure globale ($r = 0,69$, $p < 0,0001$), démontrant ainsi la présence d'une bonne validité de convergence. Aussi, le fait que la mesure multi-items de satisfaction corrélait plus fortement avec la mesure globale de satisfaction des visiteurs qu'avec les mesures de concepts différents démontre la présence de validité de construit. Il en est de même pour la mesure globale de satisfaction, puisqu'elle corrèle davantage avec la mesure multi-items qu'avec les autres concepts du modèle théorique.

Tableau 3.4 Corrélations entre les variables du modèle théorique de la satisfaction des clients lors de l'analyse d'items préliminaire ($n = 65$)

	Satisfaction (mesure globale)	Attitude	Intention de revenir	Intention de faire du bouche-à- oreille positif
Satisfaction (mesure multi- items)	0,69***	0,52***	0,40***	0,38**
Satisfaction (mesure globale)		0,44***	0,48***	0,46***
Attitude			0,42***	0,44***
Intention de revenir				0,55***

** p < 0,01 *** p < 0,001

Le tableau 3.5, quant à lui, résume les résultats des analyses de fidélité réalisées pour chacun des cinq outils de mesure. La discussion de ces résultats suit après le tableau.

Tableau 3.5 Résultats des analyses de fidélité réalisées avec les données du prétest ($n = 65$)

Outil de mesure	n ^{bre} d'items	α	Difficulté des items			Écart-type des items			Discrimination des items			n ^{bre} de corr. inter-items	
			Moy.	Min.	Max.	Moy.	Min.	Max.	Moy.	Min.	Max.	+	-
Satisfaction (mesure multi- items)	45	0,95	5,10	4,40	5,74	0,80	0,44	1,17	0,53	0,25	0,75	969	21
Satisfaction (mesure globale)	3	0,76	5,23	5,18	5,29	0,65	0,58	0,76	0,60	0,50	0,70	3	0
Attitude ⁷	13	0,87	5,24	4,75	5,52	0,90	0,66	1,25	0,59	0,28	0,80	78	0
Intention de revenir	4	0,88	3,50	0,90	5,23	1,82	0,79	3,06	0,87	0,84	0,90	6	0
Intention de faire du bouche-à-oreille positif	2	0,79	5,03	4,97	5,09	0,91	0,87	0,96	0,66	0,66	0,66	1	0

⁷ La codification de cinq des items de cet outil de mesure (items 49, 50, 52, 55 et 61) a été inversée étant donné qu'ils reflétaient une attitude défavorable, contrairement aux huit autres items qui eux reflétaient une attitude favorable.

Satisfaction des visiteurs : mesure multi-items

Les résultats de l'analyse de fidélité ont démontré que la mesure multi-items de satisfaction pouvait être utilisée telle quelle dans l'étude principale. Les indices de difficulté des items, c'est-à-dire leurs moyennes, variaient entre 4,40 et 5,74 ($M = 5,10$). L'écart-type moyen des items était de 0,80. En ce qui a trait aux indices de discrimination, ou corrélations item-total corrigées, ils variaient entre 0,25 et 0,75 avec une moyenne de 0,53, ce qui est très acceptable. De même, des 990 corrélations inter-items, les 21 qui étaient négatives n'étaient pas statistiquement significatives. Enfin, cet outil de mesure possédait un très bon indice de fidélité ($\alpha = 0,95$).

Satisfaction des visiteurs : mesure globale

Pour la mesure de satisfaction globale, l'analyse de fidélité a démontré que cet instrument de mesure pouvait aussi être utilisé tel quel dans l'étude principale. Les indices de difficulté des items, c'est-à-dire leurs moyennes, variaient entre 5,18 et 5,29 ($M = 5,23$). L'écart-type moyen des items était quant à lui de 0,65. Les indices de discrimination, ou corrélations items-total corrigées, étaient très acceptables; ils variaient entre 0,50 et 0,70 et leur valeur moyenne était de 0,60. Quant aux corrélations entre les items, elles étaient toutes positives. Sa fidélité, estimée à l'aide du coefficient alpha de Cronbach, était de 0,76. Elle était donc acceptable, d'autant plus qu'elle ne comprend que trois items.

Attitude envers la fréquentation du Musée de la civilisation

Les résultats de l'analyse de fidélité ont montré que cette mesure d'attitude était adéquate, comme le montre les données présentées au tableau 3.5. Trois des 13 items ont néanmoins

été supprimés sur la base des corrélations inter-items. En effet, trois paires d'items corrélaient de manière relativement élevée à cause de contenus trop similaires. La première paire était composée des items suivants : « Pour moi, aller au Musée de la civilisation est une activité ennuyante » (item 49) et « Pour moi, aller au Musée de la civilisation est une activité trop sérieuse » (item 50) ($r = 0,74$, $p < 0,0001$). Le premier des deux items a été conservé au détriment de l'autre étant donné que son indice de discrimination était plus élevé (0,62 contre 0,56). La deuxième paire comprenait les items 53 et 54, soit respectivement « Pour moi, aller au Musée de la civilisation est un bon choix de sortie » et « Pour moi, aller au Musée de la civilisation est du temps de loisir bien investi » ($r = 0,72$, $p < 0,0001$). Le deuxième item a été retenu puisque son indice de discrimination était plus élevé que celui du premier item (0,77 contre 0,71). La dernière paire d'items était composée des items « Pour moi, aller au Musée de la civilisation me permet de voir des choses intéressantes » (item 58) et « Pour moi, aller au Musée de la civilisation me permet d'apprendre des choses intéressantes » (item 59) ($r = 0,84$, $p < 0,0001$). Toujours sur la base de la valeur de l'indice de discrimination, le premier des deux items a été conservé dans la version finale de l'échelle d'attitude puisque son indice était de 0,73 comparativement à 0,64.

Une fois ces trois items supprimés, les différents indices psychométriques calculés lors de l'analyse de fidélité sont demeurés comparables. Les indices de difficulté des items oscillaient toujours entre 4,75 et 5,52 ($M = 5,25$) et l'écart-type moyen des items est demeuré à 0,90. Les indices de discrimination des items avaient des valeurs minimale, maximale et moyenne qui étaient respectivement de 0,31, 0,78 et de 0,55. En ce qui a trait aux corrélations entre les items, elles sont toutes restées positives. Finalement, l'indice de fidélité a quelque peu diminué à cause des items retirés (0,82 c. 0,87), mais était encore tout à fait acceptable.

Intention de revenir au Musée de la civilisation

Les résultats de l'analyse de fidélité présentés au tableau 3.5 montrent que la mesure d'intention de revenir au Musée de la civilisation était adéquate. Toutefois, lors de cette analyse, les items « J'ai l'intention de revenir au Musée de la civilisation dans les six prochains mois » (item 64) et « Je tiens fortement à revenir au Musée de la civilisation dans les six prochains mois » (item 65) sont apparus trop similaires ($r = 0,90, p < 0,0001$) pour être conservés tous les deux. L'item 65 (Je tiens fortement à revenir au Musée de la civilisation dans les six prochains mois) a par conséquent été retiré de la mesure, malgré que son indice de discrimination était légèrement supérieur à celui du premier item (0,90 comparativement à 0,88). Cette décision a été motivée par les résultats des entrevues cognitives qui avaient révélé que cet item était perçu comme étant trop extrême par trois des huit participants.

La nouvelle version de cet instrument de mesure comportait donc trois items et les indices psychométriques calculés pour ce dernier lors de l'analyse de fidélité ont peu varié. Les indices de difficulté allaient toujours de 0,90 à 5,23 ($M = 3,42$) et l'écart-type moyen des items (1,85) était également comparable. De même, les indices de discrimination avaient des valeurs minimale et maximale de 0,83 et 0,86 avec une moyenne située à 0,84. Les six corrélations inter-items sont restées positives. Enfin, le coefficient de fidélité alpha de Cronbach a diminué à cause du retrait de l'item redondant, passant de 0,88 à 0,79.

Intention de faire du bouche-à-oreille positif

Les résultats de l'analyse de fidélité démontraient que cette mesure était également acceptable. Les indices de difficulté de ses deux items étaient de 4,97 et de 5,09 ($M = 5,03$) et leurs écarts-types de 0,87 et de 0,96 ($M = 0,91$), respectivement. La corrélation entre les

deux items de l'échelle était de 0,66 ($p < 0,0001$) et le coefficient alpha de Cronbach estimait sa fidélité à 0,79.

Plan d'analyse des données

Avant de terminer ce chapitre et de poursuivre avec la présentation des résultats de l'étude, nous décrivons ici le plan d'analyse des données, c'est-à-dire la stratégie retenue pour mettre à l'épreuve nos 15 hypothèses de recherche.

Tout d'abord, rappelons que les hypothèses de recherche ont été vérifiées uniquement auprès de la mesure de satisfaction de type multi-items, et ce, principalement parce que c'est cette opérationnalisation qui est la plus courante dans la littérature scientifique et dans les entreprises. D'autres raisons expliquent également ce choix. Premièrement, vérifier les 15 hypothèses de recherche auprès des deux mesures de satisfaction aurait fait doubler le nombre déjà grand d'analyses à réaliser. Deuxièmement, la mesure de satisfaction de type globale compte trop peu d'unités d'analyse (*i.e.*, trois items) pour faire l'objet de tests statistiques.

Par la suite, il faut également mentionner que les réponses des participants aux items de satisfaction ont été converties dans une même unité de mesure de manière à ce qu'elles soient comparables peu importe qu'elles aient été exprimées sur une échelle de réponse en six ou en dix points⁸. Pour ce faire, nous avons utilisé la méthode employée dans les études similaires, c'est-à-dire de convertir les réponses en pourcentage à l'aide de la formule

⁸ Cette transformation mathématique a été appliquée aux réponses des participants aux items des deux mesures de satisfaction (multi-items et globale).

mathématique ci-après (équation 1) (Cummins, 1995, 2003; Davern & Cummins, 2006; Preston & Colman, 2000; Yu, 2003). Selon cette formule, le minimum et le maximum d'une échelle de réponse correspondent toujours à « 0 » et « 100 », respectivement. Si l'échelle de réponse a un point milieu, ce dernier obtient le score de « 50 ». À titre d'exemple, dans notre étude, la réponse « 9 » sur une échelle de réponse en dix points devenait « 88,89 » et la réponse « 5 » sur une échelle de réponse en six points devenait « 80 ».

$$P_i = (S_{ij} - 1 / R_j - 1) \times 100 \quad (\text{Éq. 1})$$

où P_i = score en pourcentage de l'individu i ;

S_{ij} = score de l'individu i sur l'échelle de réponse j ;

R_j = nombre de points dans l'échelle de réponse j .

Quant au plan d'analyse des données, il a dû tenir compte du fait que notre devis de recherche est un plan factoriel incomplet (voir le tableau 3.6). En effet, rappelons-le, nous n'avons pu croiser complètement les facteurs « nombre de points de l'échelle de réponse » et « étiquetage de l'échelle de réponse » à cause de la difficulté de développer des étiquettes pour chacun des choix de réponse de l'échelle en dix points. En conséquence, aucune des deux échelles de réponse en dix points ne comportait d'étiquettes (cellules X_{31} et X_{41} du tableau 3.6), alors que deux des quatre échelles de réponse en six points étaient étiquetées (cellules X_{12} et X_{22} du tableau 3.6). Concrètement, ceci signifie qu'il était impossible d'analyser si l'étiquetage de l'échelle de réponse modère l'effet de la polarité dans le cas des échelles de réponse en dix points. De même, il était inapproprié d'inclure les échelles de réponse en six points avec des étiquettes dans les analyses évaluant l'effet du nombre de points de l'échelle de réponse. En conséquence, les effets du nombre de points et de l'étiquetage de l'échelle de réponse ne pouvaient être analysés dans un même modèle statistique, par exemple à l'aide d'une ANOVA à trois facteurs avec effets d'interaction,

car les estimations des paramètres de ces deux facteurs n'auraient pas été indépendantes l'une de l'autre.

Tableau 3.6 Plan factoriel $2 \times 2 \times 2$ incomplet

Nombre de points	Polarité	Étiquetage des points	
		Non	Oui
6	Bipolaire	X ₁₁	X ₁₂
	Unipolaire	X ₂₁	X ₂₂
10	Bipolaire	X ₃₁	-
	Unipolaire	X ₄₁	-

En raison de notre devis de recherche, chacune des hypothèses relatives à l'effet de la polarité de l'échelle de réponse (H_1 à H_9) a été vérifiée séparément pour les échelles de réponse en six points (cellules X₁₁, X₁₂, X₂₁ et X₂₂ du tableau 3.6) et pour celles en dix points (cellules X₃₁ et X₄₁ du tableau 3.6). Pour les échelles de réponse en six points, les variables indépendantes sont la polarité et l'étiquetage de l'échelle de réponse, de manière à vérifier si l'étiquetage de l'échelle de réponse peut modérer l'effet de la polarité. Pour les échelles de réponse en dix points, seul l'effet de la polarité a été testé. Quant aux hypothèses relatives à l'effet du nombre de points de l'échelle de réponse (H_{10} à H_{15}), elles ont été vérifiées uniquement pour les échelles de réponse sans étiquettes (cellules X₁₁, X₂₁, X₃₁ et X₄₁ du tableau 3.6). Pour ces analyses, la seule variable indépendante est le nombre de points de l'échelle de réponse.

Afin de vérifier ces multiples hypothèses de recherche, plusieurs tests statistiques différents ont été nécessaires. Ceux-ci ont été réalisés à l'aide du logiciel SAS version 9.2, sauf avis contraire. Plus précisément, l'analyse de variance⁹ a été utilisée pour vérifier l'effet de la polarité de l'échelle de réponse sur les indices de difficulté (H_1), les écarts-types (H_2), les coefficients d'asymétrie (H_3) et les indices de discrimination (H_6) des items de satisfaction, ainsi que sur la moyenne du score global de satisfaction (H_4). De même, elle a été employée

⁹ Les analyses de variance ont été réalisées à l'aide de la procédure GLM du logiciel SAS.

pour vérifier l'effet du nombre de points de l'échelle de réponse sur les coefficients de variation (H_{10}) et les indices de discrimination (H_{12}) des items de satisfaction. Ensuite, le test d'homogénéité des variances de Brown-Forsythe¹⁰ a été utilisé pour vérifier l'effet de la polarité (H_5) et du nombre de points (H_{11}) de l'échelle de réponse sur la variance du score global de satisfaction. Le test développé par Hakstian et Whalen (1976) a servi, quant à lui, à vérifier si le coefficient alpha de Cronbach varie en fonction de la polarité (H_7) et du nombre de points (H_{13}) de l'échelle de réponse. Ce test a été réalisé à partir de la macrocommande Microsoft Excel développée à cet effet (Suen, 2009). Après, comme notre indice de validité de convergence correspond au coefficient de corrélation entre les deux mesures de satisfaction, la transformation r à z de Fisher a été employée de manière à être en mesure de tester si cet indice varie en fonction de la polarité (H_8) et du nombre de points (H_{14}) de l'échelle de réponse. Enfin, nous avons modélisé à l'aide d'un système d'équations structurelles les liens théoriques entre la satisfaction des clients, l'attitude envers le comportement et les intentions comportementales (voir la figure 3.7 à la page suivante) dans le but d'estimer la validité nomologique de la mesure de satisfaction. Plus particulièrement, ce sont les trois paramètres suivants qui ont servi d'indices de validité nomologique : (1) la force de l'association, ou covariance, entre la satisfaction et l'attitude envers le comportement, (2) le coefficient structurel entre la satisfaction et les intentions comportementales ainsi que (3) l'erreur résiduelle, traduisant le pouvoir prédictif du modèle. Ainsi, afin d'examiner si la validité nomologique de la mesure de satisfaction varie en fonction de la polarité (H_9) et du nombre de points (H_{15}) de l'échelle de réponse, nous avons testé l'invariance de ces trois paramètres en fonction des niveaux de ces deux variables indépendantes. Ces analyses ont été menées avec le logiciel EQS version 6.1.

¹⁰ Le test d'homogénéité des variances développé par Brown et Forsythe a été privilégié étant donné qu'il est relativement robuste aux déviations par rapport à la distribution normale (Glass & Hopkins, 1996).

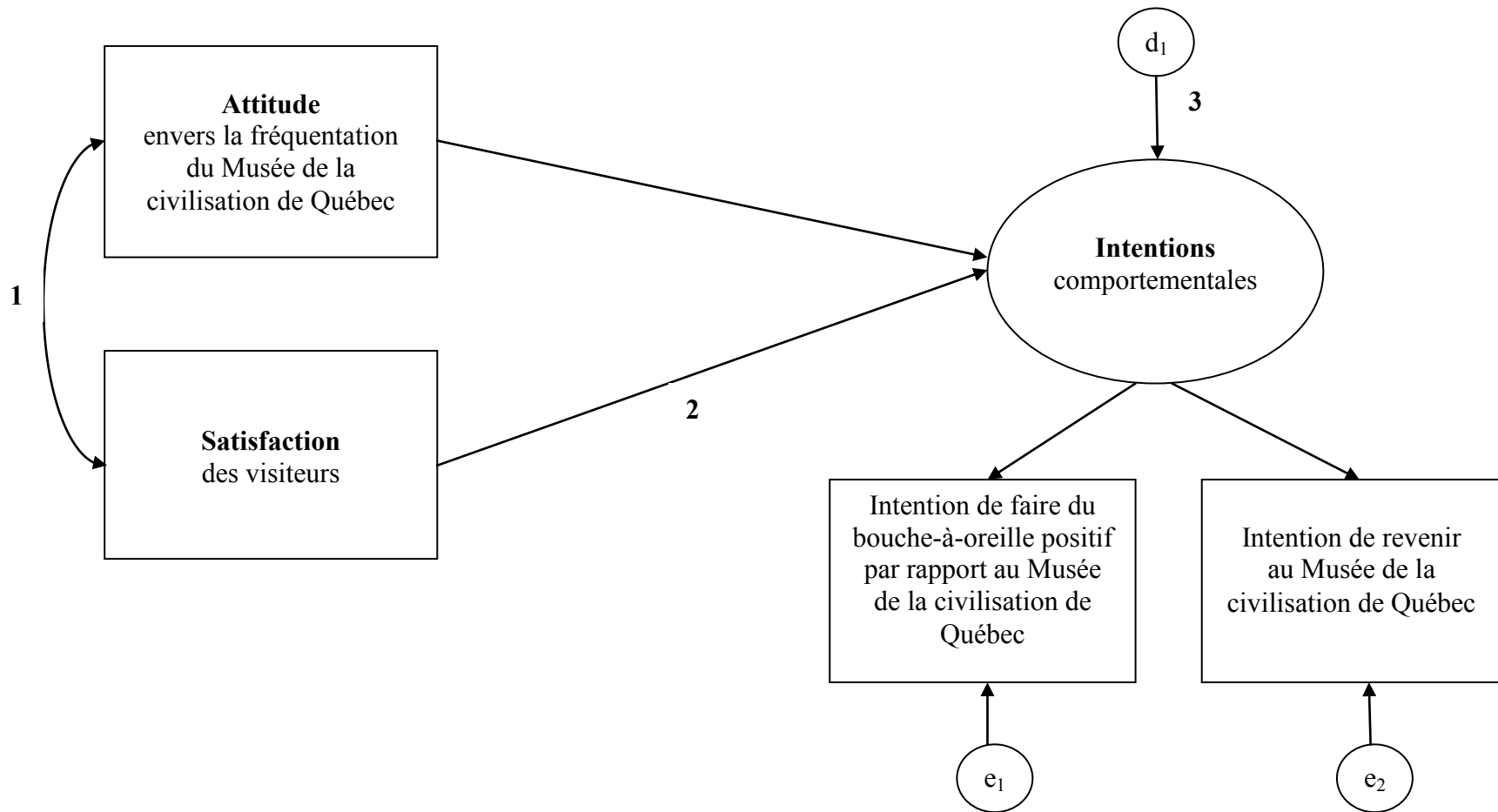


Figure 3.7 Validité nomologique : modèle structurel des relations théoriques entre la satisfaction des clients, l'attitude envers le comportement et les intentions comportementales

Avant de terminer, précisons que deux matrices de données ont été utilisées pour vérifier nos hypothèses de recherche, selon que l'hypothèse testée se rapporte aux caractéristiques des items de satisfaction (p. ex. leurs moyennes, écarts-types, coefficients d'asymétrie, indices de discrimination, etc.) ou aux caractéristiques du questionnaire de satisfaction (p. ex. sa moyenne, sa variance, son coefficient alpha de Cronbach, sa validité de convergence, etc.). Plus précisément, pour les hypothèses où les variables dépendantes sont des caractéristiques des items de satisfaction (H_1 , H_2 , H_3 , H_6 , H_{10} et H_{12}), la matrice de données analysée considère les items de satisfaction comme les « sujets » de l'analyse (lignes de la matrice), alors que les variables en colonnes consignent, pour chaque item de satisfaction, les valeurs des variables dépendantes (p. ex. les indices de difficulté, les écarts-types) et des variables indépendantes étudiées (*i.e.*, polarité, nombre de points et étiquetage de l'échelle de réponse). Par exemple, pour vérifier si les indices de discrimination des items de satisfaction varient selon le nombre de points dans l'échelle de réponse (H_6), la matrice de données analysée considère les items de satisfaction comme les « sujets » de l'analyse (lignes de la matrice), leurs indices de discrimination comme la variable dépendante (en colonne) et le nombre de points dans l'échelle de réponse constituent la variable indépendante (également en colonne). De cette manière, les indices de discrimination des 90 items de satisfaction dont l'échelle de réponse comprend dix points (cellules X_{31} et X_{41} du tableau 3.6) sont comparés à ceux des 90 items de satisfaction dont l'échelle de réponse compte six points (sans étiquettes) (cellules X_{11} et X_{21} du tableau 3.6). Ensuite, pour les hypothèses où les variables dépendantes sont des caractéristiques du questionnaire de satisfaction (H_4 , H_5 , H_7 , H_8 , H_9 , H_{11} , H_{13} , H_{14} et H_{15}), la matrice de données est de type traditionnel, c'est-à-dire que les lignes correspondent aux répondants et que les colonnes contiennent leurs réponses aux items du questionnaire en plus d'indiquer à quel groupe de comparaison appartient chacun des répondants. Par exemple, pour vérifier si le coefficient alpha de Cronbach du questionnaire de satisfaction varie en fonction du nombre de points de l'échelle de réponse (H_{13}), cette matrice de données est utilisée pour estimer la valeur de ce coefficient de fidélité selon que l'échelle de réponse compte six ou dix points.

Pour terminer, rappelons que le chapitre suivant présente les résultats des analyses statistiques réalisées pour vérifier les 15 hypothèses de recherche. Ces résultats sont présentés séparément pour chacune des hypothèses de recherche.

Chapitre 4 : Résultats

Ce quatrième chapitre présente les résultats de l'étude. Il décrit d'abord les caractéristiques psychométriques des cinq outils de mesure développés. Il enchaîne par la suite avec la mise à l'épreuve des 15 hypothèses de recherche. Cette vérification débute par les hypothèses de recherche concernant l'effet de la polarité de l'échelle de réponse, soit H_1 à H_9 , et se termine par celles relatives à l'effet du nombre de points de l'échelle de réponse, c'est-à-dire H_{10} à H_{15} .

Caractéristiques psychométriques des outils de mesure utilisés dans l'étude

Rappelons que les réponses des participants aux items de satisfaction ont été converties en pourcentages de manière à ce qu'elles soient comparables peu importe qu'elles aient été rapportées sur une échelle de réponse en six ou en dix points. Ainsi, sauf avis contraire, les réponses des participants à ces items peuvent varier entre 0 et 100. La formule mathématique utilisée pour cette transformation linéaire est celle qui a été présentée au chapitre précédent (équation 1, p. 107).

Nous décrivons ici les caractéristiques psychométriques des cinq outils de mesure utilisés dans l'étude ($n = 907$). Le tableau 4.1 qui suit présente la moyenne, l'écart-type et les valeurs minimale et maximale du score global, le coefficient de fidélité (alpha de Cronbach), les indices de discrimination des items et les corrélations inter-items pour les mesures multi-items et globale de satisfaction, et ce, de manière séparée pour chacune des six conditions testées. Le tableau 4.2, quant à lui, présente les mêmes informations, mais pour les outils de mesure de l'attitude envers la fréquentation du Musée de la civilisation,

de l'intention de revenir à ce musée et de l'intention de faire du bouche-à-oreille positif concernant ce musée.

À la lecture de ces deux tableaux, il apparaît que les différents outils de mesure ont des indices de fidélité tout à fait acceptables, ceux-ci étant supérieurs au seuil minimal de 0,70 généralement recommandé (Nunnally, 1978). Plus précisément, le coefficient alpha de Cronbach varie entre 0,77 (adéquat) et 0,98 (excellent). De même, les corrélations inter-items sont positives et les indices de discrimination des items sont supérieurs à 0,25. Les scores globaux sont relativement élevés, ce qui signifie que les items de tous les outils de mesure sont plutôt faciles à endosser pour les individus de la population étudiée. Ce constat est fort probablement attribuable au fait que l'échantillon ne provient pas de la population générale, mais plutôt d'une sous-population, soit celle des visiteurs du Musée de la civilisation. Cette sous-population risque davantage d'apprécier les expositions du Musée de la civilisation, d'entretenir une attitude favorable à l'égard de sa fréquentation et d'avoir des intentions comportementales favorables envers ce musée.

Tableau 4.1 Score global et indices psychométriques des deux outils de mesure de la satisfaction des clients pour chacune des six conditions testées

Mesure	n	α	Score global				Corr. item-total			Corr. inter-items		
			Moy.	É.-T.	Min.	Max.	Moy.	Min.	Max.	Moy.	Min.	Max.
Multi-items (45 items)												
Condition 1	141	0,97	79,48	12,78	8,89	100,00	0,68	0,49	0,80	0,47	0,17	0,86
Condition 2	152	0,97	73,01	11,97	16,00	100,00	0,64	0,34	0,80	0,42	0,05	0,86
Condition 3	154	0,98	75,27	15,91	2,67	99,11	0,72	0,49	0,83	0,53	0,15	0,89
Condition 4	153	0,96	76,82	11,82	44,44	100,00	0,60	0,37	0,78	0,96	0,02	0,78
Condition 5	157	0,98	78,54	13,36	14,32	100,00	0,71	0,55	0,83	0,51	0,21	0,90
Condition 6	150	0,98	77,06	13,86	7,41	98,97	0,74	0,46	0,86	0,56	0,22	0,91
Globale (3 items)												
Condition 1	141	0,91	80,05	17,55	0,00	100,00	0,83	0,77	0,87	0,79	0,72	0,87
Condition 2	152	0,88	73,75	17,45	6,67	100,00	0,77	0,69	0,84	0,71	0,62	0,82
Condition 3	154	0,90	76,93	18,13	6,67	100,00	0,80	0,74	0,85	0,75	0,69	0,84
Condition 4	153	0,84	75,58	16,74	20,00	100,00	0,71	0,65	0,79	0,65	0,56	0,74
Condition 5	157	0,92	78,87	16,87	11,11	100,00	0,84	0,76	0,88	0,79	0,73	0,91
Condition 6	150	0,90	77,17	16,69	0,00	100,00	0,80	0,73	0,84	0,75	0,69	0,85

Note. Condition 1 = échelle de réponse en six points bipolaire (avec étiquettes)
 Condition 2 = échelle de réponse en six points unipolaire (avec étiquettes)
 Condition 3 = échelle de réponse en six points bipolaire (sans étiquettes)
 Condition 4 = échelle de réponse en six points unipolaire (sans étiquettes)
 Condition 5 = échelle de réponse en dix points bipolaire (sans étiquettes)
 Condition 6 = échelle de réponse en dix points unipolaire (sans étiquettes)

Tableau 4.2 Score global et indices psychométriques des outils de mesure de l'attitude envers la fréquentation du musée et des intentions comportementales

Mesure	n	Items	α	Score global				Corr. item-total			Corr. inter-items		
				Moy.	É.-T.	Min.	Max.	Moy.	Min.	Max.	Moy.	Min.	Max.
Attitude	907	10	0,83	5,09	0,69	1,60	6,00	0,54	0,27	0,71	0,35	0,13	0,79
Intention de revenir	907	3	0,77	11,46	5,50	2,00	26,00	0,74	0,60	0,81	0,65	0,56	0,81
Intention de faire du b.-à-oreille positif	907	2	0,84	4,84	1,01	2,00	6,00	0,73	0,73	0,73	0,73	0,73	0,73

Note. Les scores de quatre des items de l'outil de mesure de l'attitude envers la fréquentation du Musée de la civilisation (items 49, 51, 53 et 58) ont été inversés étant donné qu'ils reflétaient une attitude défavorable, contrairement aux huit autres items qui, eux, reflétaient une attitude favorable.

Mise à l'épreuve des hypothèses de recherche

Cette partie du quatrième chapitre présente les analyses réalisées pour éprouver les hypothèses de recherche. Le premier bloc d'analyses vise à vérifier les hypothèses relatives à l'effet de la polarité de l'échelle de réponse, soit H_1 à H_9 , et le second s'intéresse aux hypothèses concernant le nombre de points de l'échelle de réponse, soit H_{10} à H_{15} .

Hypothèses 1 à 9 : effet de la polarité de l'échelle de réponse

Pour débiter, rappelons que les résultats des études présentées dans le cadre théorique ne permettaient pas de déduire avec conviction que le fait d'étiqueter ou pas les points de l'échelle de réponse interagissait avec la polarité sur les réponses de satisfaction. Pour cette raison, ce facteur a été intégré au devis de recherche à titre exploratoire, sans émettre d'hypothèse formelle à son sujet. Toutefois, à cause de la grande difficulté à créer des étiquettes pour une échelle de réponse en dix points, ce facteur a été croisé avec la variable « polarité de l'échelle de réponse » uniquement pour les échelles de réponse en six points. En conséquence, à l'intérieur de ce premier bloc d'analyses, chaque hypothèse est d'abord vérifiée pour les échelles de réponses en six points, en tenant compte de l'étiquetage de l'échelle de réponse, et ensuite pour celles en dix points, en évaluant uniquement l'effet de la polarité.

H_1 : les items de satisfaction dont l'échelle de réponse est unipolaire présentent des indices de difficulté inférieurs à ceux des items dont l'échelle de réponse est bipolaire

Échelles de réponse en six points

Cette première hypothèse de recherche a d'abord été mise à l'épreuve auprès des items de satisfaction dont l'échelle de réponse compte six points. Le tableau 4.3 présente les valeurs moyennes des indices de difficulté des items de satisfaction selon la polarité et l'étiquetage de l'échelle de réponse.

Afin de tester l'effet de la polarité sur les indices de difficulté des items de satisfaction et de vérifier si l'étiquetage de l'échelle de réponse peut modérer cet effet, une ANOVA à deux facteurs et un terme d'interaction a été réalisée. Les résultats de cette analyse (tableau 4.4) indiquent que ces deux facteurs et leur interaction expliquent 26 % de la variance des indices de difficulté des items de satisfaction et que seuls l'effet d'interaction et l'effet simple de la polarité sont significatifs au seuil alpha de 0,05. En conséquence, l'effet de la polarité sur les indices de difficulté des items de satisfaction a été examiné en fonction de cette interaction. Cet effet d'interaction est illustré à la figure 4.1. Deux analyses de contrastes ont été réalisées afin de vérifier si l'effet de la polarité sur les indices de difficulté des items de satisfaction varie selon que l'échelle de réponse possède ou pas des étiquettes. Le risque de commettre une erreur de type 1 a été contrôlé à l'aide de la correction de Bonferonni, de sorte que le seuil de signification a été fixé à 0,025. Les résultats de ces analyses de contrastes (tableau 4.5) démontrent que l'effet de la polarité est effectivement modéré par l'étiquetage de l'échelle de réponse, celui-ci étant significatif uniquement lorsque l'échelle de réponse possède des étiquettes. Dans cette condition, la polarité de l'échelle de réponse explique 24 % de la variance des indices de difficulté des items de satisfaction, ce qui correspond à une grande taille d'effet ($\omega^2 = 0,24 > 0,14$)

d'après les normes suggérées par Cohen (1988)¹¹. Toujours dans cette situation et allant dans le sens de H_1 , les items de satisfaction dont l'échelle de réponse est unipolaire plutôt que bipolaire ont des indices de difficulté significativement inférieurs (73,04 c. 79,49). Ainsi, à partir de ces résultats, H_1 est acceptée dans le contexte des échelles de réponse en six points, mais uniquement lorsque ces dernières possèdent des étiquettes.

Tableau 4.3 Valeurs moyennes des indices de difficulté des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points)

Facteur	Niveau	Indice de difficulté
Polarité (P)	Bipolaire	77,38
	Unipolaire	74,93
Étiquettes (E)	Oui	76,26
	Non	76,04

Tableau 4.4 ANOVA testant l'effet de la polarité et de l'étiquetage sur les indices de difficulté des items de satisfaction (échelles de réponse en six points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	ω^2	R^2
Modèle	3	20,84*	< 0,0001		0,26
Polarité (P)	1	17,03*	< 0,0001	0,07	
Étiquettes (E)	1	0,14	0,7093	0,00	
P × E	1	45,36*	< 0,0001	0,19	
Erreur	176				

* $p < 0,05$

¹¹ Kirk (1995) suggère d'employer les normes proposées par Cohen (1988) pour interpréter une taille d'effet à partir de la statistique oméga carré. Selon ces normes, des valeurs de 0,01, 0,06 et 0,14 représentent respectivement des effets de taille petite, moyenne et grande.

Tableau 4.5 Analyses de contrastes évaluant l'effet de la polarité sur les indices de difficulté des items de satisfaction selon l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	Indice de difficulté	$F(1, 176)$	p	ω^2
Oui	Bipolaire	79,49	58,99*	< 0,0001	0,24
	Unipolaire	73,04			
Non	Bipolaire	75,27	3,40	0,0668	0,01
	Unipolaire	76,82			

* $p < 0,025$

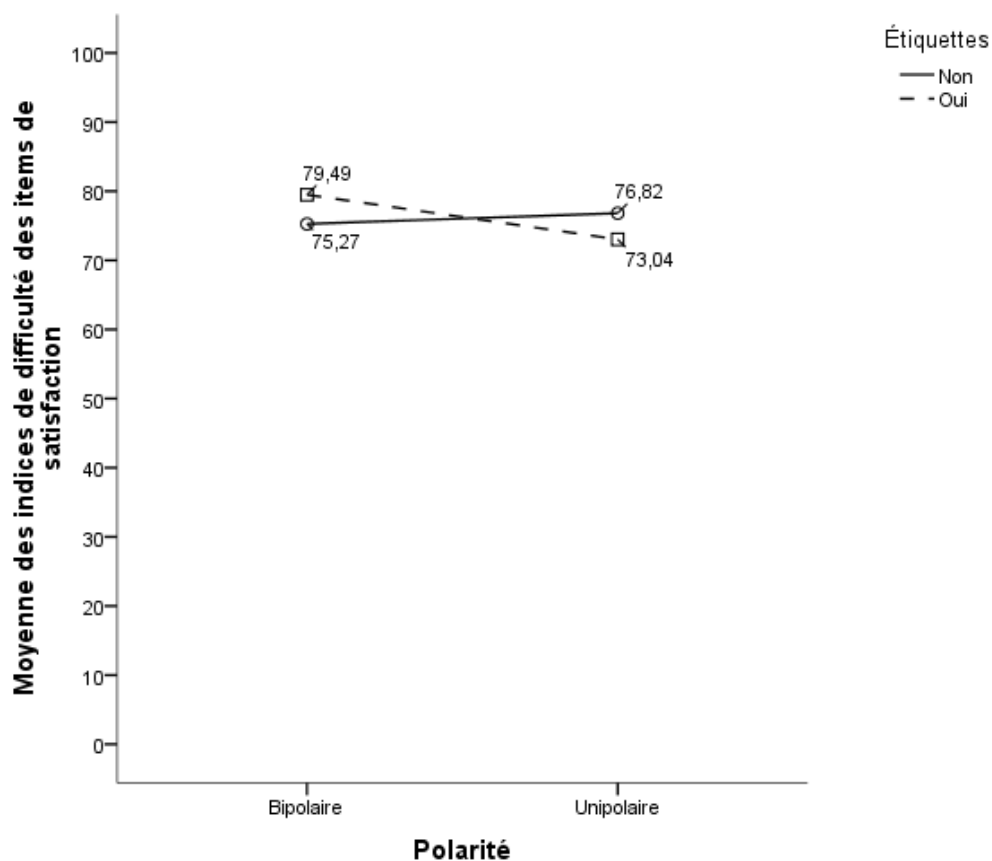


Figure 4.1 Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les indices de difficulté des items de satisfaction (échelles de réponse en six points)

Échelles de réponse en dix points

La première hypothèse de recherche a ensuite été mise à l'épreuve auprès des items de satisfaction dont l'échelle de réponse comprend dix points. Dans cette condition, les indices de difficultés des items de satisfaction ont une valeur moyenne de 78,56 ($E.-T. = 3,47$) lorsque l'échelle de réponse est bipolaire et de 77,03 ($E.-T. = 3,84$) lorsque cette dernière est unipolaire. Une ANOVA à un facteur a été réalisée afin de tester l'effet de la polarité de l'échelle de réponse sur les indices de difficulté des items de satisfaction. Les résultats de cette ANOVA (tableau 4.6) indiquent que cet effet est non significatif et donc que H_1 ne peut être acceptée dans le cas où l'échelle de réponse comprend dix points.

Tableau 4.6 ANOVA testant l'effet de la polarité sur les indices de difficulté des items de satisfaction (échelles de réponse en dix points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Modèle	1	3,91	0,0511	0,04
Erreur	88			

* $p < 0,05$

H_2 : les items de satisfaction dont l'échelle de réponse est unipolaire présentent des écarts-types supérieurs aux items dont l'échelle de réponse est bipolaire

Échelles de réponse en six points

Le tableau 4.7 présente les valeurs moyennes des écarts-types des items de satisfaction selon la polarité et l'étiquetage de l'échelle de réponse. Afin de tester l'effet de la polarité sur les écarts-types des items de satisfaction et de vérifier si l'étiquetage de l'échelle de

réponse peut modérer cet effet, une ANOVA à deux facteurs avec effet d'interaction a été réalisée (tableau 4.8). Les résultats de cette analyse indiquent que ces deux facteurs et leur interaction expliquent 34 % de la variance des écarts-types des items de satisfaction et que les deux effets simples et leur interaction sont significatifs au seuil alpha de 0,05.

L'effet de la polarité sur les écarts-types des items de satisfaction a donc été examiné en fonction de cette interaction (figure 4.2). Deux analyses de contrastes ont été réalisées pour tester l'effet de la polarité selon que l'échelle de réponse est étiquetée ou pas. Le risque de commettre une erreur de type 1 a été contrôlé à l'aide de la correction de Bonferonni de sorte que le seuil de signification a été fixé à 0,025. Les résultats de ces analyses de contrastes (tableau 4.9) démontrent que l'effet de la polarité sur les écarts-types des items de satisfaction est effectivement modéré par l'étiquetage de l'échelle de réponse, cet effet étant significatif uniquement lorsque l'échelle de réponse ne possède pas d'étiquettes. La polarité de l'échelle de réponse explique 20 % de la variance des écarts-types des items de satisfaction, ce qui correspond à une grande taille d'effet ($\omega^2 = 0,20 > 0,14$) d'après les normes suggérées par Cohen (1988). Toutefois, à l'encontre de la deuxième hypothèse de recherche, les écarts-types des items de satisfaction sont inférieurs lorsque l'échelle de réponse est unipolaire (19,07) plutôt que bipolaire (21,70). Ainsi, à partir de ces résultats, il apparaît qu'on ne peut accepter H_2 dans le contexte des échelles de réponse en six points.

Tableau 4.7 Valeurs moyennes des écarts-types des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points)

Facteur	Niveau	É.-T.
Polarité (P)	Bipolaire	20,37
	Unipolaire	18,87
Étiquettes (E)	Oui	18,84
	Non	20,39

Tableau 4.8 ANOVA testant l'effet de la polarité et de l'étiquetage sur les écarts-types des items de satisfaction (échelles de réponse en six points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	ω^2	R^2
Modèle	3	30,48*	< 0,0001*		0,34
Polarité (P)	1	35,00*	< 0,0001*	0,13	
Étiquettes (E)	1	36,75*	< 0,0001*	0,13	
P × E	1	19,69*	< 0,0001*	0,07	
Erreur	176				

* $p < 0,05$

Tableau 4.9 Analyses de contrastes évaluant l'effet de la polarité sur les écarts-types des items de satisfaction selon l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	<i>É.-T.</i>	<i>F</i> (1, 176)	<i>p</i>	ω^2
Oui	Bipolaire	19,04	1,09	0,2971	0,00
	Unipolaire	18,66			
Non	Bipolaire	21,70	53,59*	< 0,0001	0,20
	Unipolaire	19,07			

* $p < 0,025$

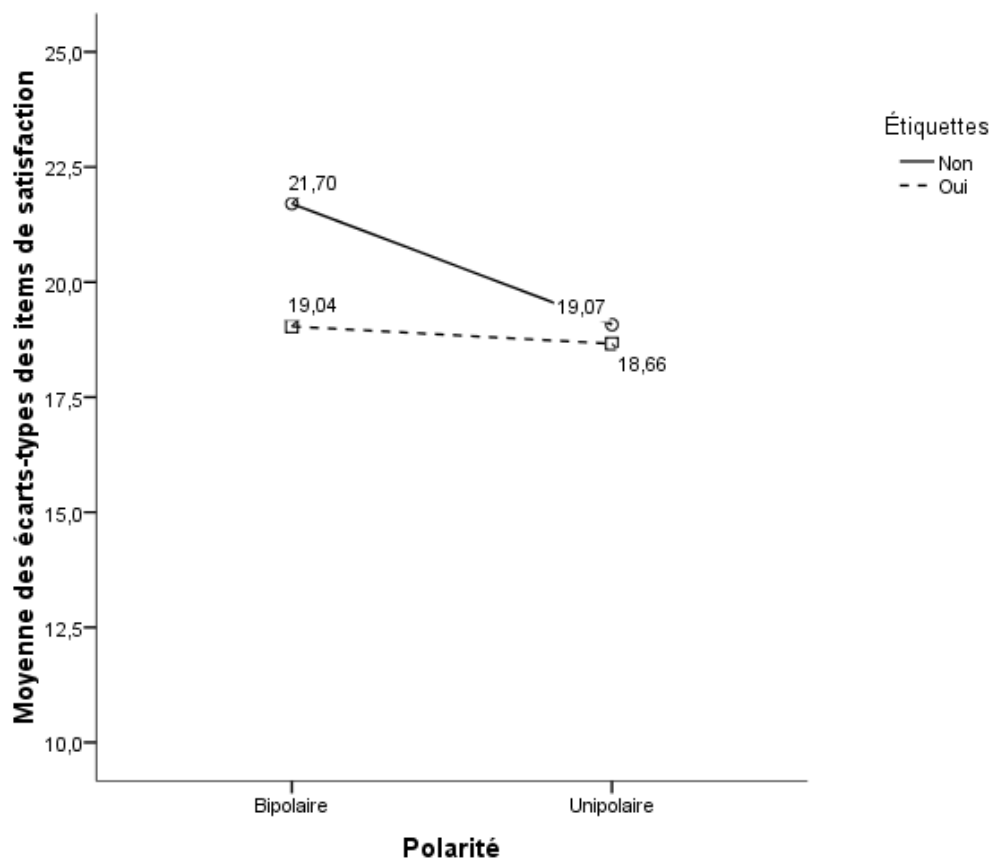


Figure 4.2 Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les écarts-types des items de satisfaction (échelles de réponse en six points)

Échelles de réponse en dix points

La deuxième hypothèse de recherche a ensuite été mise à l'épreuve pour les items de satisfaction dont l'échelle de réponse compte dix points. Pour ces items de satisfaction, l'écart-type moyen est de 18,56 lorsque l'échelle de réponse est bipolaire et de 19,01 lorsqu'elle est unipolaire. Une ANOVA à un facteur a été réalisée afin de tester l'effet de la polarité de l'échelle de réponse sur les écarts-types des items de satisfaction. Les résultats

de cette ANOVA (tableau 4.10) indiquent que cet effet est non significatif et ne permettent d'accepter H_2 dans le cas où l'échelle de réponse comprend dix points.

Tableau 4.10 ANOVA testant l'effet de la polarité sur les écarts-types des items de satisfaction (échelles de réponse en dix points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	R^2
Modèle	1	1,45	0,2325	0,02
Erreur	88			

* $p < 0,05$

H_3 : les items de satisfaction dont l'échelle de réponse est unipolaire présentent des indices d'asymétrie moins importants que les items dont l'échelle de réponse est bipolaire

Échelles de réponse en six points

Le tableau 4.11 présente les moyennes des indices d'asymétrie des items de satisfaction selon la polarité et l'étiquetage de l'échelle de réponse. Une ANOVA à deux facteurs avec effet d'interaction a été menée pour mettre à l'épreuve l'effet de la polarité et de l'étiquetage sur les indices d'asymétrie des items de satisfaction. Les résultats de cette analyse (tableau 4.12) indiquent que le modèle testé explique 32 % de la variance des indices d'asymétrie des items de satisfaction et que seuls l'effet d'interaction et l'effet simple de la polarité sont significatifs au seuil alpha de 0,05. Ainsi, l'effet de la polarité sur les indices d'asymétrie des items de satisfaction a été examiné en fonction de cette interaction (voir la figure 4.3). Deux analyses de contrastes ont été réalisées afin de vérifier si l'effet de la polarité varie selon que l'échelle de réponse possède ou pas des étiquettes (tableau 4.13). La correction de Bonferonni a été utilisée pour contrôler le risque de

commettre une erreur de type 1 de sorte que le seuil de signification a été fixé à 0,025. Les résultats de ces analyses de contrastes démontrent que l'effet de la polarité sur les indices d'asymétrie des items de satisfaction est modéré par l'étiquetage de l'échelle de réponse. Par ailleurs, bien que cet effet soit significatif avec ou sans étiquetage, il est beaucoup plus important lorsque l'échelle de réponse est étiquetée. Plus spécifiquement, dans la situation où l'échelle de réponse possède des étiquettes, la polarité explique 26 % de la variance des indices d'asymétrie des items de satisfaction, ce qui correspond à une grande taille d'effet ($\omega^2 = 0,26 > 0,14$) d'après les normes suggérées par Cohen (1988). En comparaison, lorsque l'échelle de réponse ne possède pas d'étiquettes, la polarité n'explique que 6 % de la variance des indices d'asymétrie des items de satisfaction, ce qui ne représente qu'un effet modéré ($\omega^2 = 0,06 \geq 0,06$). Néanmoins, l'utilisation d'une échelle de réponse unipolaire est associée à des indices d'asymétrie des items de satisfaction significativement inférieurs que l'échelle de réponse soit étiquetée (-0,65 c. -1,30) ou pas (-0,77 c. -1,07). Conséquemment, H_3 est acceptée dans le contexte des échelles de réponse en six points.

Tableau 4.11 Valeurs moyennes des indices d'asymétrie des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points)

Facteur	Niveau	Asymétrie
Polarité (P)	Bipolaire	-1,19
	Unipolaire	-0,71
Étiquettes (E)	Oui	-0,98
	Non	-0,92

Tableau 4.12 ANOVA testant l'effet de la polarité et de l'étiquetage sur les indices d'asymétrie des items de satisfaction (échelles de réponse en six points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	ω^2	R^2
Modèle	3	28,20*	< 0,0001		0,32
Polarité (P)	1	74,18*	< 0,0001	0,28	
Étiquettes (E)	1	1,03	0,3121	0,00	
P × E	1	9,40*	0,0025	0,03	
Erreur	176				

* $p < 0,05$

Tableau 4.13 Analyses de contrastes évaluant l'effet de la polarité sur les indices d'asymétrie des items de satisfaction selon l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	Asymétrie	$F(1, 176)$	<i>p</i>	ω^2
Oui	Bipolaire	-1,30	68,20*	< 0,0001	0,26
	Unipolaire	-0,65			
Non	Bipolaire	-1,07	15,38*	0,0001	0,06
	Unipolaire	-0,77			

* $p < 0,025$

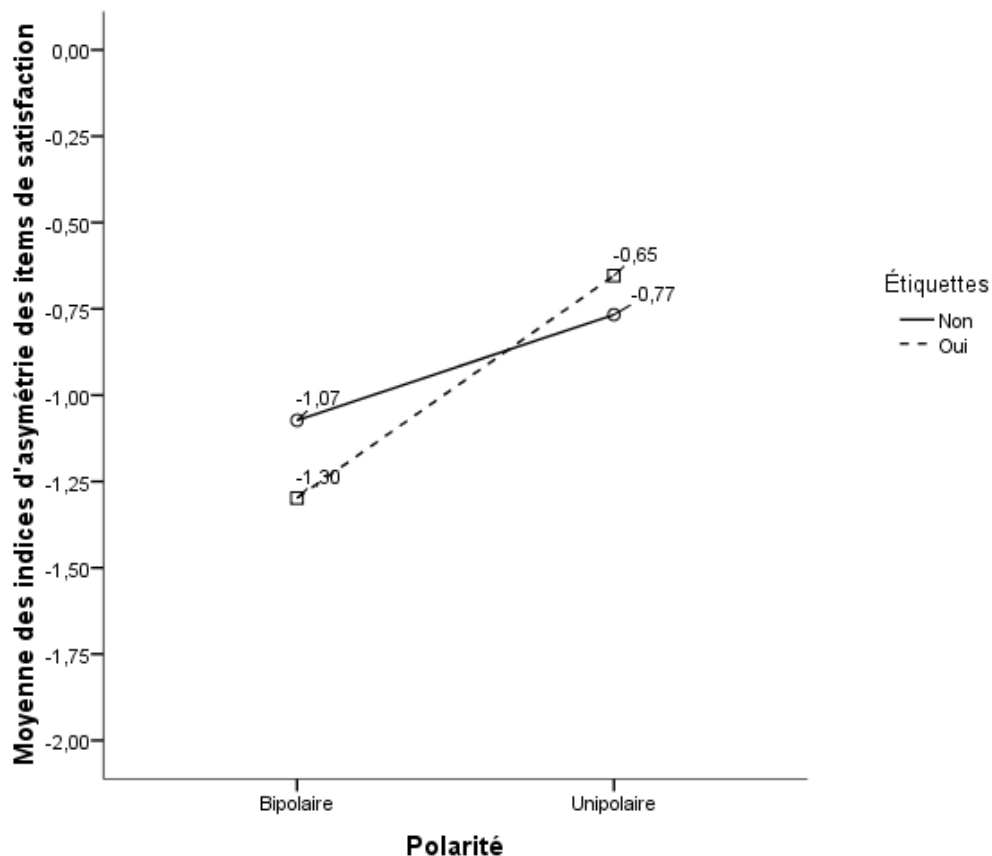


Figure 4.3 Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les indices d'asymétrie des items de satisfaction (échelles de réponse en six points)

Échelles de réponse en dix points

La troisième hypothèse de recherche a ensuite été mise à l'épreuve auprès des items de satisfaction dont l'échelle de réponse comprend dix points. Une ANOVA à un facteur a été menée pour vérifier si les indices d'asymétrie des items de satisfaction diffèrent de manière significative selon que l'échelle de réponse est bipolaire ou unipolaire. Les résultats de l'ANOVA (tableau 4.14) indiquent que le degré d'asymétrie des items de satisfaction n'est pas moins important lorsque l'échelle de réponse est unipolaire ($M = -1,32$, $É.-T. = 0,44$)

plutôt que bipolaire ($M = -1,31$, $\acute{E}.-T. = 0,48$), ce qui ne nous permet pas d'accepter H_3 pour les échelles de réponse en dix points.

Tableau 4.14 ANOVA testant l'effet de la polarité sur les indices d'asymétrie des items de satisfaction (échelles de réponse en dix points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	R^2
Modèle	1	0,02	0,8790	0,00
Erreur	88			

* $p < 0,05$

H_4 : la satisfaction exprimée dans un questionnaire dont l'échelle de réponse est unipolaire est moins élevée que celle exprimée dans un questionnaire dont l'échelle de réponse est bipolaire

Échelles de réponse en six points

Le tableau 4.15 présente le score global moyen au questionnaire de satisfaction selon la polarité et l'étiquetage de l'échelle de réponse. Afin de tester l'effet de la polarité sur le score global de satisfaction et de vérifier si l'étiquetage de l'échelle de réponse peut modérer cet effet, une ANOVA à deux facteurs avec effet d'interaction a été réalisée. Les résultats de cette analyse (tableau 4.16) indiquent que ce modèle statistique est significatif au seuil alpha de 0,05, mais qu'il n'explique que 3 % de la variance du score global de satisfaction, l'effet d'interaction et l'effet simple de la polarité étant significatifs. Cet effet d'interaction est illustré à la figure 4.4. Deux analyses de contrastes ont été réalisées pour vérifier si l'effet de la polarité sur le score global de satisfaction diffère selon que l'échelle de réponse possède ou pas des étiquettes (tableau 4.17). La correction de Bonferroni a été appliquée de manière à contrôler le risque de commettre une erreur de type 1 et le seuil de signification a ainsi été fixé à 0,025. Les résultats de ces analyses de contrastes démontrent

que l'effet de la polarité est effectivement modéré par l'étiquetage de l'échelle de réponse, car celui-ci est significatif uniquement lorsque l'échelle de réponse possède des étiquettes. Dans cette condition, la polarité de l'échelle de réponse explique 3 % de la variance du score global de satisfaction, ce qui correspond à une petite taille d'effet ($0,01 < \omega^2 = 0,03 < 0,06$) d'après les normes suggérées par Cohen (1988). Néanmoins, cet effet significatif va dans le sens de la quatrième hypothèse de recherche, c'est-à-dire que l'emploi d'une échelle de réponse unipolaire est associé à un score moyen de satisfaction inférieur (73,01 c. 79,48). En conséquence, H_4 est acceptée dans le contexte des échelles de réponse en six points, mais uniquement lorsque ces dernières possèdent des étiquettes.

Tableau 4.15 Score global de satisfaction en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)

Facteur	Niveau	<i>n</i>	Moyenne
Polarité (P)	Bipolaire	295	77,28
	Unipolaire	305	74,92
Étiquettes (E)	Oui	293	76,12
	Non	307	76,04

Tableau 4.16 ANOVA testant l'effet de la polarité et de l'étiquetage sur le score global de satisfaction (échelles de réponse en six points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	ω^2	R^2
Modèle	3	6,18*	0,0004		0,03
Polarité (P)	1	5,19*	0,0231	0,01	
Étiquettes (E)	1	0,03	0,8570	0,00	
P × E	1	13,73*	0,0002	0,02	
Erreur	596				

* $p < 0,05$

Tableau 4.17 Analyses de contrastes évaluant l'effet de la polarité sur le score global de satisfaction selon l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	<i>n</i>	Moyenne	<i>F</i> (1, 596)	<i>p</i>	ω^2
Oui	Bipolaire	141	79,48	17,48*	< 0,0001	0,03
	Unipolaire	152	73,01			
Non	Bipolaire	154	75,27	1,04	0,3073	0,00
	Unipolaire	153	76,82			

* $p < 0,025$

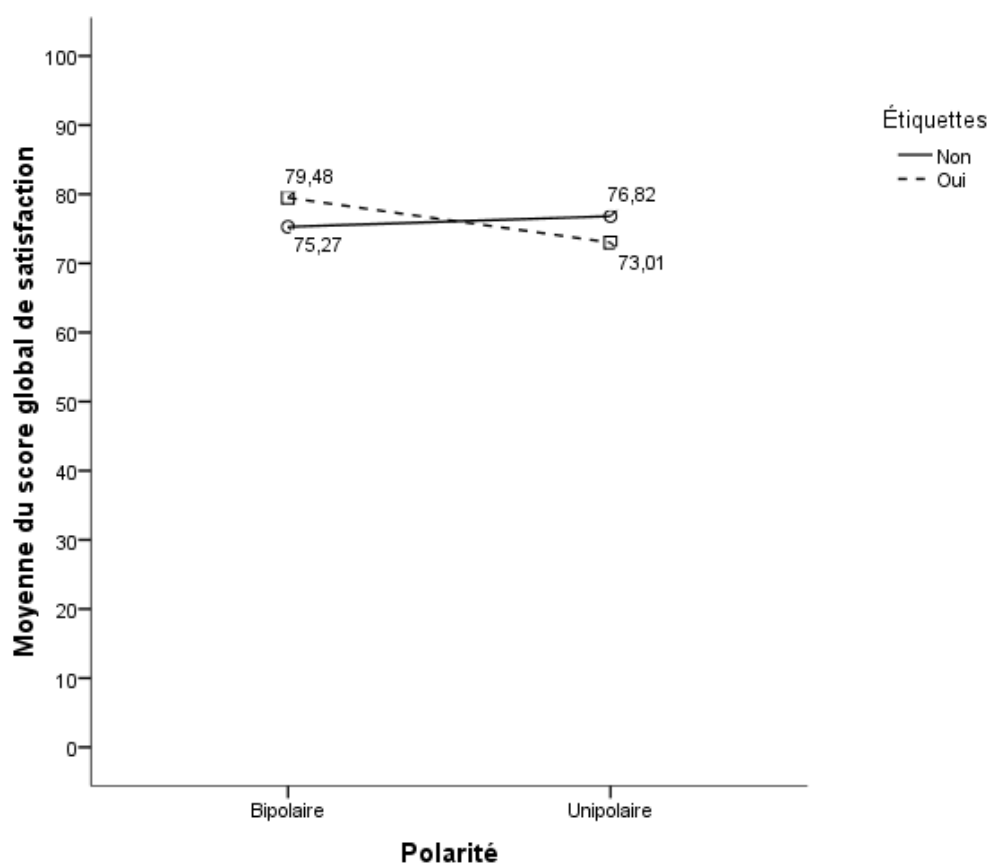


Figure 4.4 Illustration de l'effet d'interaction entre la polarité et l'étiquetage des points intermédiaires de l'échelle de réponse sur le score global de satisfaction (échelles de réponse en six points)

Échelles de réponse en dix points

La quatrième hypothèse de recherche a ensuite été mise à l'épreuve pour les versions du questionnaire de satisfaction dont l'échelle de réponse comprend dix points. Les résultats indiquent que le score global moyen de satisfaction est de 78,54 quand l'échelle de réponse est bipolaire et de 77,06 lorsqu'elle est unipolaire. Une ANOVA à un facteur a été réalisée afin de tester l'effet de la polarité de l'échelle de réponse sur le score global de satisfaction. Les résultats de cette ANOVA (tableau 4.18) indiquent que cet effet est non significatif au seuil alpha de 0,05 et, en conséquence, H_4 ne peut être acceptée dans le cas où l'échelle de réponse comporte dix points.

Tableau 4.18 ANOVA testant l'influence de la polarité sur le score global de satisfaction (échelles de réponse en dix points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Modèle	1	0,91	0,3396	0,00
Erreur	305			

* $p < 0,05$

H_5 : la variance de la satisfaction exprimée dans un questionnaire dont l'échelle de réponse est unipolaire est plus grande que celle exprimée dans un questionnaire dont l'échelle de réponse est bipolaire.

Échelles de réponse en six points

C'est le test de Brown-Forsythe¹² qui a servi à comparer la variance du score global de satisfaction en fonction de la polarité et de l'étiquetage de l'échelle de réponse. Cependant, ce test statistique ne permet pas de vérifier s'il existe un effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse. Or, les analyses précédentes réalisées pour tester H_2 ont identifié que ces deux effets interagissent sur les écarts-types des items de satisfaction. Pour tenir compte de ce résultat et pour pallier au fait que le test de Brown-Forsythe ne permet pas de tester d'effet d'interaction, nous avons décidé de vérifier l'effet de la polarité de l'échelle de réponse de manière séparée pour les échelles de réponse avec et sans étiquettes.

Le tableau 4.19 présente la variance du score global de satisfaction en fonction de la polarité de l'échelle de réponse selon que cette dernière est étiquetée ou pas de même que les résultats des tests de Brown-Forsythe. D'après ces résultats, la polarité de l'échelle de réponse n'affecte pas significativement la variance du score global de satisfaction, et ce, peu importe que l'échelle de réponse soit étiquetée ou pas. La cinquième hypothèse de recherche est donc rejetée dans le cas où l'échelle de réponse possède six points.

¹² Le test d'homogénéité des variances développé par Brown et Forsythe a été privilégié étant donné qu'il est relativement robuste aux déviations par rapport à la distribution normale (Glass & Hopkins, 1996). Le test de Levene arrive toutefois aux mêmes conclusions que ce dernier.

Tableau 4.19 Variance du score global de satisfaction en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	<i>n</i>	Variance	Brown-Forsythe	
				<i>F</i>	<i>p</i>
Oui	Bipolaire	141	163,45	$F(1,291) = 0,79$	0,3756
	Unipolaire	152	143,18		
Non	Bipolaire	154	252,98	$F(1, 305) = 1,45$	0,2300
	Unipolaire	153	139,64		

* $p < 0,05$

Échelles de réponse en dix points

Le tableau 4.20 présente la variance du score global de satisfaction selon la polarité de l'échelle de réponse lorsque cette dernière comporte dix points. Le test de Brown-Forsythe a permis de vérifier si la polarité de l'échelle de réponse a un effet sur la variance du score global de satisfaction. Le résultat de ce test (tableau 4.20) montre que la polarité de l'échelle de réponse n'a pas d'effet significatif sur la variance du score global de satisfaction. La cinquième hypothèse de recherche n'est donc également pas acceptée dans le cas où l'échelle de réponse possède dix points.

Tableau 4.20 Variance du score global de satisfaction en fonction de la polarité (échelles de réponse en dix points)

Polarité	<i>n</i>	Variance	Brown-Forsythe	
			<i>F</i>	<i>p</i>
Bipolaire	157	178,41	$F(1,305) = 0,07$	0,7848
Unipolaire	150	192,10		

* $p < 0,05$

H_6 : les indices de discrimination des items de satisfaction dont l'échelle de réponse est unipolaire sont supérieurs à ceux dont l'échelle de réponse est bipolaire

Échelles de réponse en six points

Le tableau 4.21 présente les indices de discrimination moyens des items de satisfaction selon la polarité et l'étiquetage de l'échelle de réponse. Afin de tester l'effet de la polarité sur les indices de discrimination des items de satisfaction et de vérifier si l'étiquetage de l'échelle de réponse peut modérer cet effet, une ANOVA à deux facteurs avec effet d'interaction a été réalisée. Les résultats de cette analyse (tableau 4.22) indiquent que ce modèle statistique est significatif au seuil alpha de 0,05 et qu'il explique 19 % de la variance des indices de discrimination des items de satisfaction. Toujours selon ces résultats, l'effet d'interaction (voir la figure 4.5) et l'effet simple de la polarité sont significatifs. Deux analyses de contrastes ont été réalisées afin de vérifier l'effet de la polarité selon que l'échelle de réponse possède ou pas des étiquettes (tableau 4.23). Le risque de commettre une erreur de type 1 a été contrôlé à l'aide de la correction de Bonferonni et le seuil de signification a été fixé à 0,025. Les résultats de ces analyses de contrastes démontrent que l'effet de la polarité est effectivement modéré par l'étiquetage de l'échelle de réponse, celui-ci étant significatif seulement si l'échelle de réponse ne possède pas d'étiquettes. Dans cette situation, la polarité de l'échelle de réponse explique 17 % de la variance des indices de discrimination des items de satisfaction, ce qui correspond à une grande taille d'effet ($\omega^2 = 0,17 > 0,14$) d'après les normes suggérées par Cohen (1988). Cependant, à l'opposé de la sixième hypothèse de recherche, les indices de discrimination des items de satisfaction sont inférieurs lorsque l'échelle de réponse est unipolaire (0,60 c. 0,72). La sixième hypothèse de recherche ne peut donc être acceptée dans le contexte des échelles de réponse en six points.

Tableau 4.21 Moyennes des indices de discrimination des items de satisfaction selon la polarité et l'étiquetage (échelles de réponse en six points)

Facteur	Niveau	Indice de discrimination moyen
Polarité (P)	Bipolaire	0,70
	Unipolaire	0,62
Étiquettes (E)	Oui	0,66
	Non	0,66

Tableau 4.22 ANOVA testant l'effet de la polarité et de l'étiquetage sur les indices de discrimination des items de satisfaction (échelles de réponse en six points)

	<i>d.l.</i>	<i>F</i>	<i>p</i>	ω^2	R^2
Modèle	3	14,14*	< 0,0001*		0,19
Polarité (P)	1	33,06	< 0,0001*	0,15	
Étiquettes (E)	1	0,00	0,9786	0,00	
P × E	1	9,37*	0,0026*	0,04	
Erreur	176				

* $p < 0,05$

Note. Les variances des quatre groupes résultant du croisement entre les deux variables indépendantes ne sont pas homogènes selon le test de Brown-Forsythe, $F(3, 176) = 3,63, p = 0,0140$. Néanmoins, le test F de Welch indique qu'au moins un des quatre groupes diffère significativement, $F(3, 96,28) = 13,63, p < 0,0001$.

Tableau 4.23 Analyses de contrastes évaluant l'effet de la polarité sur les indices de discrimination des items de satisfaction selon l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	Indice de discrimination moyen	$F(1, 176)$	<i>p</i>	ω^2
Oui	Bipolaire	0,68	3,62	0,0588	0,01
	Unipolaire	0,64			
Non	Bipolaire	0,72	38,81*	< 0,0001	0,17
	Unipolaire	0,60			

* $p < 0,025$

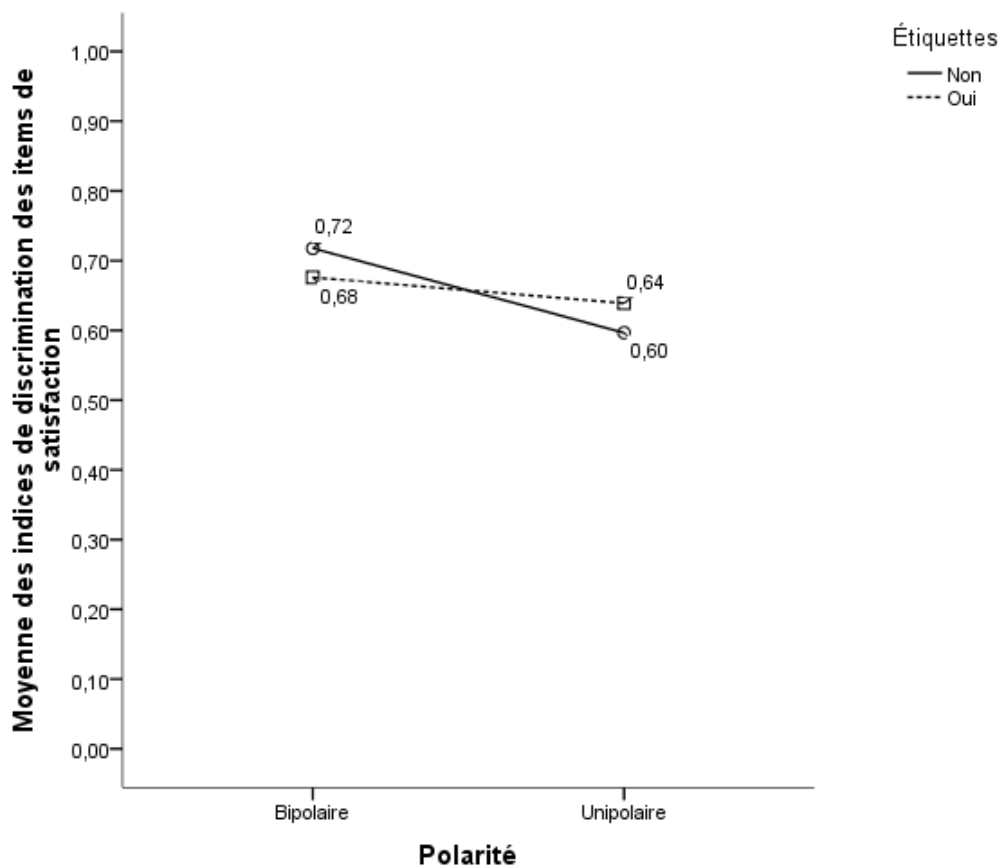


Figure 4.5 Illustration de l'effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse sur les indices de discrimination des items de satisfaction (échelles de réponse en six points)

Échelles de réponse en dix points

La sixième hypothèse de recherche a ensuite été mise à l'épreuve auprès des items de satisfaction pour lesquels l'échelle de réponse comprend dix points. Une ANOVA à un facteur a été réalisée afin de tester l'effet de la polarité de l'échelle de réponse sur les indices de discrimination des items de satisfaction. Les résultats de cette ANOVA (tableau 4.24) indiquent que les indices de discrimination des items de satisfaction ne sont pas supérieurs quand l'échelle de réponse est unipolaire ($M = 0,74$, $E.-T. = 0,08$) plutôt que

bipolaire ($M = 0,71$, $E.-T. = 0,07$). En conséquence, H_6 ne peut être acceptée dans le cas où l'échelle de réponse comprend dix points.

Tableau 4.24 ANOVA testant l'effet de la polarité sur les indices de discrimination des items de satisfaction (échelles de réponse en dix points)

Source	<i>d.l.</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Modèle	1	3,80	0,0543	0,04
Erreur	176			

* $p < 0,05$

H_7 : le coefficient alpha de Cronbach d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celui d'un questionnaire dont l'échelle de réponse est bipolaire

Échelles de réponse en six points

Le test de Hakstien et Whalen (1976) a été utilisé pour comparer les valeurs du coefficient alpha de Cronbach en fonction de la polarité et de l'étiquetage de l'échelle de réponse du questionnaire de satisfaction. Toutefois, ce test ne permet pas de vérifier s'il existe un effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse du questionnaire de satisfaction. Comme les analyses pour vérifier H_6 ont identifié une interaction entre ces deux variables sur les indices de discrimination des items de satisfaction, nous avons décidé de tester l'effet de la polarité de l'échelle de réponse sur le coefficient alpha de Cronbach de manière indépendante pour les échelles de réponse avec et sans étiquettes.

Le tableau 4.25 présente les valeurs du coefficient alpha de Cronbach pour le questionnaire de satisfaction en fonction de la polarité de l'échelle de réponse selon que cette dernière est étiquetée ou pas, de même que les résultats des tests de Hakstian et Whalen (1976). Les résultats indiquent que la polarité de l'échelle de réponse a un effet sur le coefficient alpha de Cronbach, mais uniquement dans le cas où celle-ci ne possède pas d'étiquettes. Contrairement à notre septième hypothèse de recherche, l'utilisation d'une échelle de réponse unipolaire est associée à un coefficient alpha de Cronbach légèrement inférieur (0,9629 c. 0,9798) au plan statistique.

Tableau 4.25 Coefficients alpha de Cronbach selon la polarité et de l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	n	α	Hakstian et Whalen	
				$\chi^2_{(df=1)}$	p
Oui	Bipolaire	120	0,9744	0,9330	0,3341
	Unipolaire	132	0,9695		
Non	Bipolaire	133	0,9798	11,52*	0,0007
	Unipolaire	130	0,9629		

* $p < 0,05$

Échelles de réponse en dix points

Le tableau 4.26 présente les valeurs du coefficient alpha de Cronbach pour le questionnaire de satisfaction selon la polarité de l'échelle de réponse lorsque cette dernière compte dix points. Le test de Hakstian et Whalen a été utilisé pour vérifier si le coefficient alpha de Cronbach du questionnaire de satisfaction diffère de manière statistiquement significative selon la polarité de l'échelle de réponse. Le résultat de ce test (tableau 4.26) montre que la polarité de l'échelle de réponse n'a pas d'effet significatif sur le coefficient alpha de Cronbach du questionnaire de satisfaction. La septième hypothèse de recherche ne peut donc être acceptée dans le cas où l'échelle de réponse possède dix points.

Tableau 4.26 Coefficients alpha de Cronbach selon la polarité de l'échelle de réponse (échelles de réponse en dix points)

Polarité	<i>n</i>	α	Hakstian et Whalen	
			$\chi^2_{(df=1)}$	<i>p</i>
Bipolaire	130	0,9785	1,0135	0,3141
Unipolaire	120	0,9821		

* $p < 0,05$

H_8 : le degré de validité de convergence d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celui d'un questionnaire dont l'échelle de réponse est bipolaire

Échelles de réponse en six points

La validité de convergence du questionnaire de satisfaction a été estimée, rappelons-le, à l'aide du coefficient de corrélation entre les mesures multi-items et globale de satisfaction. Afin de comparer la magnitude de cette corrélation en fonction de la polarité et de l'étiquetage de l'échelle de réponse du questionnaire de satisfaction, le test de comparaison basé sur la transformation r à Z de Fisher pour les différences entre deux ou plusieurs coefficients de corrélations indépendants a été employé. Toutefois, ce test ne permet pas de vérifier la présence d'un effet d'interaction entre deux variables et les analyses réalisées précédemment pour tester H_6 et H_7 ont montré que l'effet de la polarité de l'échelle de réponse sur les indices de discrimination des items de satisfaction et sur le coefficient alpha de Cronbach du questionnaire de satisfaction varie en fonction de l'étiquetage de l'échelle de réponse. Pour cette raison, nous avons décidé de vérifier l'effet de la polarité de l'échelle de réponse sur la validité de convergence du questionnaire de satisfaction de manière séparée pour les échelles de réponse avec et sans étiquettes.

Le tableau 4.27 présente les valeurs du coefficient de corrélation entre les mesures multi-items et globale de satisfaction en fonction la polarité de l'échelle de réponse selon que celle-ci est étiquetée ou pas, d'une part, de même que les résultats des tests statistiques, d'autre part. Les résultats indiquent que la polarité de l'échelle de réponse n'a pas d'effet significatif sur le degré de validité de convergence du questionnaire de satisfaction, ce qui amène à ne pas pouvoir accepter H_8 dans la condition où des échelles de réponse en six points sont utilisées.

Tableau 4.27 Indices de validité de convergence en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	n	r^a (indice de validité de convergence)	r à z de Fisher	
				$\chi^2_{(dl=1)}$	p
Oui	Bipolaire	141	0,76	1,19	0,2751
	Unipolaire	152	0,70		
Non	Bipolaire	154	0,82	2,54	0,1107
	Unipolaire	153	0,75		

* $p < 0,05$

^a Ces coefficients de corrélation ont tous une valeur $p < 0,0001$

Échelles de réponse en dix points

Le tableau 4.28 présente le degré de validité de convergence du questionnaire de satisfaction selon la polarité de l'échelle de réponse lorsque cette dernière comporte dix points. Le test de comparaison basé sur la transformation r à Z de Fisher pour les différences entre deux ou plusieurs coefficients de corrélations indépendants a été employé pour tester si le degré de validité de convergence varie en fonction de la polarité de l'échelle de réponse. Le résultat (tableau 4.28) de ce test démontre que la polarité de l'échelle de réponse n'a pas d'effet significatif sur le degré de validité de convergence du

questionnaire de satisfaction. La huitième hypothèse de recherche ne peut donc être acceptée dans le contexte des échelles de réponse en dix points.

Tableau 4.28 Indices de validité de convergence en fonction de la polarité (échelles de réponse en dix points)

Polarité	<i>n</i>	<i>r</i> ^a (indice de validité de convergence)	<i>r</i> à <i>z</i> de Fisher	
			$\chi^2_{(df=1)}$	<i>p</i>
Bipolaire	157	0,83	1,58	0,2087
Unipolaire	150	0,87		

* $p < 0,05$

^a Ces coefficients de corrélation ont tous une valeur $p < 0,0001$

***H₉* : le degré de validité nomologique d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celle d'un questionnaire dont l'échelle de réponse est bipolaire**

Rappelons que la validité nomologique du questionnaire de satisfaction a été estimée en testant un modèle structurel correspondant au modèle théorique de la satisfaction des clients (voir la figure 3.7 du plan d'analyse des données au chapitre 3, p. 110). Plus spécifiquement, trois des paramètres de ce modèle ont servi à estimer la validité nomologique, soit (1) la force de l'association entre la satisfaction et l'attitude envers le comportement, (2) le coefficient structurel entre la satisfaction et les intentions comportementales ainsi que (3) l'erreur résiduelle, traduisant le pouvoir prédictif du modèle. L'hypothèse 9 a été vérifiée en testant l'invariance de ces trois paramètres en fonction de la polarité de l'échelle de réponse du questionnaire de satisfaction selon que cette dernière compte six ou dix points. Les prochaines sections présentent les résultats de ces tests d'abord pour la condition des échelles de réponse en six points et ensuite pour celle des échelles de réponse en dix points. Les indices d'ajustement et les coefficients de

détermination pour les différentes estimations du modèle structurel sont présentés à l'annexe 13 (tableaux A13.1 et A13.2). À titre informatif, les tableaux 4.29 et 4.30 présentent les corrélations entre la satisfaction des clients et les variables associées (*i.e.*, attitude envers la fréquentation du musée, intentions comportementales) selon la polarité de l'échelle de réponse du questionnaire de satisfaction, et ce, pour les échelles de réponse en six et en dix points, respectivement.

Tableau 4.29 Corrélations entre la satisfaction des clients et les variables associées en fonction de la polarité et de l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	<i>n</i>	Attitude	Intention de revenir	Intention de faire du bouche-à-oreille positif
Oui	Bipolaire	141	0,27**	0,22**	0,45***
	Unipolaire	152	0,50***	0,26***	0,43***
Non	Bipolaire	154	0,42***	0,13	0,37***
	Unipolaire	153	0,60***	0,27***	0,55***

* $p < 0,05$ ** $p < 0,01$ *** $p < 0,001$

Tableau 4.30 Corrélations entre la satisfaction des clients et les variables associées en fonction de la polarité (échelles de réponse en dix points)

	Polarité	<i>n</i>	Attitude	Intention de revenir	Intention de faire du bouche-à-oreille positif
	Bipolaire	157	0,23**	0,25**	0,35***
	Unipolaire	150	0,32***	0,25**	0,33***

* $p < 0,05$ ** $p < 0,01$ *** $p < 0,001$

Échelles de réponse en six points

Le tableau 4.31 donne les valeurs des trois indices de validité nomologique du questionnaire de satisfaction en fonction de la polarité de l'échelle de réponse selon que cette dernière possède ou pas des étiquettes. Étant donné que les analyses précédentes ont montré que l'étiquetage de l'échelle de réponse peut modérer l'effet de la polarité, les tests d'invariance ont été réalisés séparément pour les questionnaires de satisfaction dont l'échelle de réponse est étiquetée et pour ceux dont l'échelle de réponse n'est pas étiquetée. De plus, étant donné que trois tests d'invariance ont été nécessaires pour évaluer si la polarité de l'échelle de réponse a un effet sur la validité nomologique, soit un test pour chacun des trois indices, la correction de Bonferonni a été appliquée pour contrôler la probabilité de commettre une erreur de type 1. Le seuil de signification statistique a été ainsi corrigé à la baisse et fixé à 0,017 (0,05/3).

Les résultats des tests d'invariance des trois paramètres utilisés comme indices de validité nomologique sont présentés au tableau 4.32. Il appert qu'aucune des valeurs du chi-carré n'est significative au seuil alpha de 0,017. Selon ces résultats, le degré de validité nomologique du questionnaire de satisfaction ne varie pas de manière significative en fonction de la polarité de l'échelle de réponse. En conséquence, H_0 ne peut être acceptée pour les questionnaires de satisfaction ayant une échelle de réponse en six points.

Tableau 4.31 Indices de validité nomologique en fonction de la polarité selon l'étiquetage (échelles de réponse en six points)

Étiquettes	Polarité	<i>n</i>	Satisfaction – attitude	Satisfaction – intentions comportementales	<i>d</i> ₁
Oui	Bipolaire	141	0,27 (2,34)	0,38 (0,03)	0,78 (1,00)
	Unipolaire	152	0,50 (4,35)	0,34 (0,02)	0,79 (1,00)
Non	Bipolaire	154	0,42 (4,33)	0,25 (0,02)	0,89 (1,00)
	Unipolaire	153	0,60 (4,67)	0,40 (0,03)	0,65 (1,00)

Note. Les coefficients sont tous significatifs à $p < 0,05$. Les valeurs présentées sont les coefficients standardisés et, entre parenthèses, les coefficients non standardisés.

Tableau 4.32 Tests de l'invariance des trois paramètres structurels en fonction de la polarité selon l'étiquetage (échelles de réponse en six points)

Étiquettes	Satisfaction - attitude		Satisfaction - intentions comportementales		<i>d</i> ₁	
	χ^2	<i>p</i>	χ^2	<i>p</i>	χ^2	<i>p</i>
Oui	3,255	0,071	0,619	0,431	0,193	0,660
Non	0,082	0,775	4,109	0,043	5,221	0,022

* $p < 0,017$

Échelles de réponse en dix points

Le tableau 4.33 donne les valeurs des trois indices de validité nomologique du questionnaire de satisfaction en fonction de la polarité de l'échelle de réponse pour les questionnaires de satisfaction dont l'échelle de réponse comporte dix points. Comme dans le cas des échelles de réponses en six points, la correction de Bonferonni a été appliquée pour contrôler la probabilité de commettre une erreur de type 1 étant donné que trois tests d'invariance étaient nécessaires pour évaluer si la polarité de l'échelle de réponse a un effet sur la validité nomologique. Le seuil de signification statistique est donc également fixé à 0,017 (0,05/3).

Les résultats des tests d'invariance des trois paramètres utilisés comme indices de validité nomologique sont présentés au tableau 4.34. L'examen de ce tableau permet de constater qu'aucune des valeurs du chi-carré n'est significative au seuil alpha de 0,017. Selon ces résultats, le degré de validité nomologique du questionnaire de satisfaction ne varie pas de manière significative en fonction de la polarité de l'échelle de réponse. La neuvième hypothèse de recherche ne peut également pas être acceptée pour les questionnaires de satisfaction dont l'échelle de réponse comporte dix points.

Tableau 4.33 Indices de validité nomologique en fonction de la polarité (échelles de réponse en dix points)

Polarité	<i>n</i>	Satisfaction - attitude	Satisfaction - intentions comportementales	<i>d</i> ₁
Bipolaire	157	0,23 (2,06)	0,35 (0,02)	0,72 (1,00)
Unipolaire	150	0,32 (3,17)	0,24 (0,02)	0,78 (1,00)

Note. Les coefficients sont tous significatifs à $p < 0,05$. Les valeurs présentées sont les coefficients standardisés et, entre parenthèses, les coefficients non standardisés.

Tableau 4.34 Tests de l'invariance des trois paramètres structurels selon la polarité (échelles de réponse en dix points)

Satisfaction - attitude		Satisfaction - intentions comportementales		<i>d</i> ₁	
χ^2	<i>p</i>	χ^2	<i>p</i>	χ^2	<i>p</i>
0,978	0,323	0,921	0,337	0,659	0,417

Hypothèses 10 à 15 : effet du nombre de points dans l'échelle de réponse

Le premier bloc d'analyses a servi à tester les hypothèses H_1 à H_9 qui se rapportent à l'effet de la polarité de l'échelle de réponse. Pour chacune de ces hypothèses, l'effet de la polarité de l'échelle de réponse a été testée tant dans la condition des échelles de réponse en six points que dans la condition des échelles de réponse en dix points.

Ce deuxième bloc d'analyses, quant à lui, vise à tester les hypothèses de recherche H_{10} à H_{15} qui concernent l'effet du nombre de points dans l'échelle de réponse (six c. dix). Étant donné que nous n'avons pas construit d'échelle de réponse en dix points avec des étiquettes, seule la condition des échelles de réponse sans étiquettes a été utilisée pour tester ces hypothèses. Les tests statistiques présentés dans cette partie comparent donc les échelles de réponse en six points sans étiquettes à celles en dix points sans étiquettes.

De plus, en dépit du fait que les hypothèses H_{10} à H_{15} ne concernent que l'effet du nombre de points dans l'échelle de réponse, nous avons tout de même tenu compte de l'effet de la polarité de l'échelle de réponse dans les analyses statistiques visant à les vérifier. Cette décision a été prise parce que les résultats des analyses réalisées pour tester H_2 , H_6 et H_7 laissent penser qu'il pourrait exister un effet d'interaction entre la polarité et le nombre de points de l'échelle de réponse. En effet, ces résultats montrent que les écarts-types des items de satisfaction, leurs indices de discrimination et le coefficient alpha de Cronbach du questionnaire de satisfaction diffèrent selon la polarité de l'échelle de réponse dans la condition des échelles de réponse en six points sans étiquettes, mais pas dans celle des échelles de réponse en dix points. Ainsi, il nous est apparu pertinent de tenir compte de la polarité de l'échelle de réponse en analysant l'effet du nombre de points de l'échelle de réponse. Pour cette raison, la variable « polarité de l'échelle de réponse » figure dans les analyses présentées ci-après.

H_{10} : les items de satisfaction dont l'échelle de réponse a dix points présentent des coefficients de variation supérieurs aux items dont l'échelle de réponse a six points

Le coefficient de variation a été choisi pour comparer la variabilité des items de satisfaction en fonction du nombre de points de l'échelle de réponse. En effet, cet indice statistique permet de comparer la variabilité d'items n'ayant pas à l'origine la même unité de mesure. Le tableau 4.35 présente les valeurs moyennes des coefficients de variation des items de satisfaction selon que l'échelle de réponse comporte six ou dix points et selon qu'elle est bipolaire ou unipolaire. Une ANOVA à deux facteurs avec effet d'interaction a été réalisée afin de tester si le nombre de points et la polarité de l'échelle de réponse ont un effet sur les coefficients de variation des items de satisfaction. Les résultats de cette ANOVA (tableau 4.36) indiquent que le modèle testé est statistiquement significatif et qu'il explique 29 % de la variance des coefficients de variation des items de satisfaction. Comme l'effet d'interaction est statistiquement significatif, les deux effets simples n'ont pas été examinés davantage.

Des analyses de contrastes ont ensuite été réalisées de manière à mieux disséquer l'effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse (figure 4.6). Plus précisément, nous avons testé si l'effet du nombre de points de l'échelle de réponse sur les coefficients de variation des items de satisfaction diffère selon que l'échelle de réponse est bipolaire ou unipolaire. Comme deux analyses de contrastes étaient nécessaires pour y arriver, la probabilité de commettre une erreur de type 1 a été contrôlée à l'aide de la méthode de Bonferonni et le seuil de signification a été fixé à 0,025. Les résultats de ces analyses de contrastes sont présentés au tableau 4.37. Dans la condition des échelles de réponse bipolaires, le nombre de points de réponse explique 25 % de la variance des coefficients de variation des items de satisfaction, ce qui correspond à un effet de grande taille selon les normes de Cohen (1988) ($\omega^2 = 0,25 > 0,14$). Dans cette condition, les items de satisfaction ont des coefficients de variation inférieurs lorsque l'échelle de réponse compte dix points plutôt que six, ce qui va en sens contraire de H_{10} . Dans la condition des

échelles de réponse unipolaires, le nombre de points n'a pas d'effet sur les coefficients de variation des items de satisfaction. L'hypothèse de recherche H_{10} ne peut donc être acceptée, et ce, tant dans la condition des échelles de réponse bipolaires que dans celle des échelles de réponse unipolaires.

Tableau 4.35 Coefficients de variation moyens des items de satisfaction selon le nombre de points et la polarité de l'échelle de réponse

Facteur	Niveau	C.V.
Nombre de points (N)	6 points	26,98 %
	10 points	24,26 %
Polarité (P)	Bipolaire	26,34 %
	Unipolaire	24,90 %

Tableau 4.36 ANOVA testant l'effet du nombre de points et de la polarité de l'échelle de réponse sur les coefficients de variation des items de satisfaction

Source	d.l.	F	p	ω^2	R ²
Modèle	3	23,78*	< 0,0001		0,29
Nombre de points (N)	1	33,51*	< 0,0001	0,13	
Polarité (P)	1	9,36*	0,0026	0,03	
N × P	1	28,48*	< 0,0001	0,11	
Erreur	176				

* $p < 0,05$

Tableau 4.37 Analyses de contrastes évaluant l'effet du nombre de points sur les coefficients de variation des items de satisfaction selon la polarité de l'échelle de réponse

Polarité	Nombre de points	C.V.	$F(1, 176)$	p	ω^2
Bipolaire	6	28,94 %	61,89*	< 0,0001	0,25
	10	23,73 %			
Unipolaire	6	25,01 %	0,10	0,7496	0,00
	10	24,80 %			

* $p < 0,025$

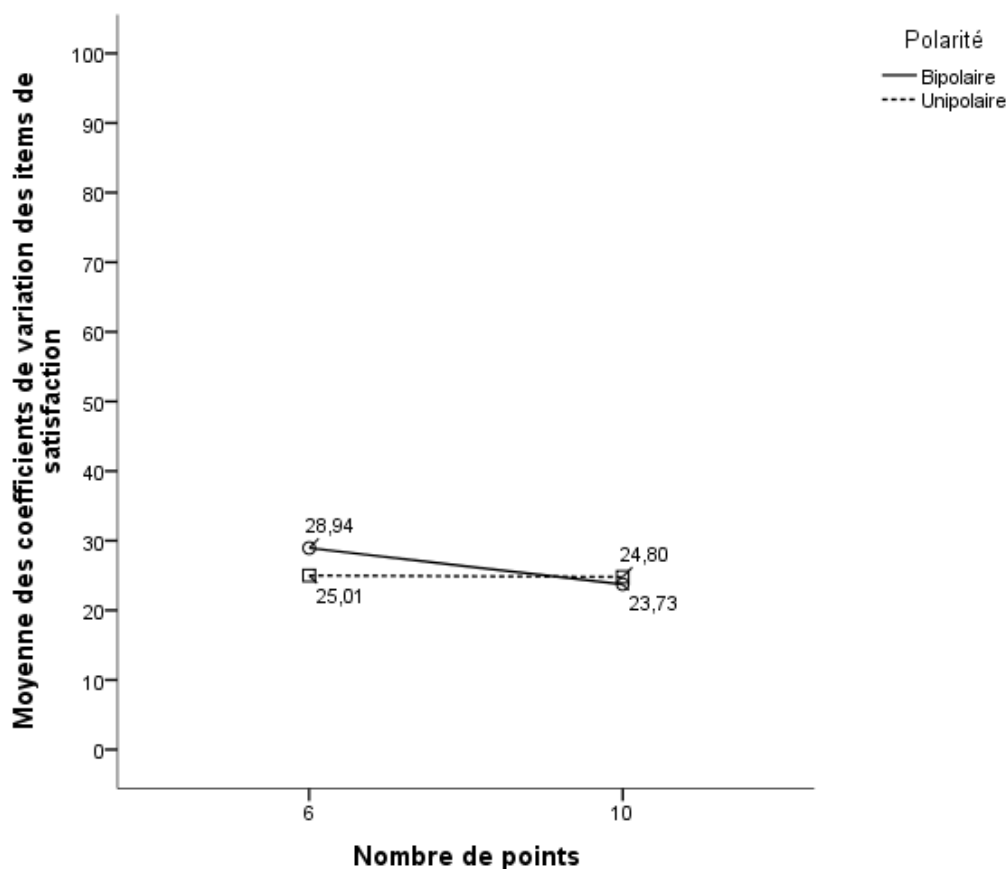


Figure 4.6 Illustration de l'effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse sur les coefficients de variation des items de satisfaction

H_{11} : la variance de la satisfaction exprimée dans un questionnaire dont l'échelle de réponses a dix points est plus élevée que celle obtenue dans un questionnaire dont l'échelle de réponse a six points

Dans le but de comparer la variabilité du score global de satisfaction en fonction du nombre de points de l'échelle de réponse, la variance calculée à partir des scores exprimés en pourcentage a été utilisée plutôt que le coefficient de variation. Effectivement, à notre connaissance, aucun test statistique ne permet de comparer directement deux coefficients de variation entre eux.

C'est le test de Brown-Forsythe qui a servi à comparer la variance du score global de satisfaction en fonction du nombre de points de l'échelle de réponse. Toutefois, ce test statistique ne permet pas de vérifier s'il existe un effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse. Or, les analyses réalisées pour tester l'hypothèse précédente (*i.e.*, H_{10}) ont identifié que les coefficients de variation des items de satisfaction varient selon le nombre de points de l'échelle de réponse, mais que cet effet diffère selon que l'échelle de réponse est bipolaire ou unipolaire. Pour tenir compte de ce résultat et pour pallier au fait que le test de Brown-Forsythe ne permet pas de tester d'effet d'interaction, nous avons décidé de vérifier l'effet du nombre de points de l'échelle de réponse de manière séparée pour les échelles de réponse bipolaires et unipolaires. Le tableau 4.38 présente la variance du score global de satisfaction en fonction du nombre de points de réponse selon que l'échelle de réponse est bipolaire ou unipolaire, d'une part, ainsi que les résultats des tests de Brown-Forsythe, d'autre part. Ces derniers montrent que le nombre de points dans l'échelle de réponse n'affecte pas de manière significative la variance du score global de satisfaction, et ce, peu importe que l'échelle de réponse soit bipolaire ou unipolaire. En conséquence, il est impossible d'accepter H_{11} .

Tableau 4.38 Variance du score global de satisfaction en fonction du nombre de points de réponse selon que l'échelle de réponse est bipolaire ou unipolaire

Polarité	Nombre de points	<i>n</i>	Variance ¹	Brown-Forsythe	
				<i>F</i>	<i>p</i>
Bipolaire	6	154	252,98	$F(1, 309) = 1,07$	0,3021
	10	157	178,41		
Unipolaire	6	153	139,64	$F(1, 301) = 0,05$	0,8270
	10	150	192,10		

* $p < 0,05$

¹Calculée à partir des scores exprimés en pourcentage.

H_{12} : les indices de discrimination des items de satisfaction dont l'échelle de réponse a dix points sont supérieurs à ceux dont l'échelle de réponse a six points

Le tableau 4.39 présente les valeurs moyennes des indices de discrimination des items de satisfaction selon le nombre de points et la polarité de l'échelle de réponse. Une ANOVA à deux facteurs avec effet d'interaction a été réalisée pour vérifier si les indices de discrimination des items de satisfaction varient selon le nombre de points et la polarité de l'échelle de réponse. Les résultats de cette analyse (tableau 4.40) indiquent que le modèle testé explique 29 % de la variance des indices de discrimination des items de satisfaction. L'effet d'interaction étant significative, les effets simples n'ont pas été examinés davantage.

Des analyses de contrastes ont par la suite été menées pour mieux comprendre l'effet d'interaction (illustrée à la figure 4.7). Plus spécifiquement, nous avons vérifié l'effet du nombre de points de l'échelle de réponse sur les indices de discrimination des items de satisfaction selon que celle-ci est bipolaire ou unipolaire. Comme deux analyses de contrastes étaient nécessaires, la correction de Bonferonni a été appliquée et le risque de

commettre une erreur de type 1 a été fixé à 0,025. Les résultats des deux analyses de contrastes réalisées sont présentés au tableau 4.41. Ces résultats indiquent que le nombre de points de réponse a un effet significatif sur les indices de discrimination des items de satisfaction, mais uniquement dans la condition où les échelles de réponse sont unipolaires. Dans cette condition, l'effet observé va dans le sens de H_{12} , c'est-à-dire que les items de satisfaction dont l'échelle de réponse a dix points plutôt que six ont des indices de discrimination supérieurs, et la taille de cet effet peut être qualifié de grande d'après les normes de Cohen (1988) ($\omega^2 = 0,24 > 0,14$). L'hypothèse 12 est donc acceptée dans la condition des échelles de réponse unipolaires, mais pas dans celle des échelles de réponse bipolaires.

Tableau 4.39 Indices de discrimination moyens en fonction du nombre de points et de la polarité de l'échelle de réponse

Facteur	Niveau	Indice de discrimination moyen
Nombre de points (N)	6 points	0,66
	10 points	0,72
Polarité (P)	Bipolaire	0,71
	Unipolaire	0,67

Tableau 4.40 ANOVA testant l'effet du nombre de points et de la polarité de l'échelle de réponse sur les indices de discrimination des items de satisfaction

	<i>d.l.</i>	<i>F</i>	<i>p</i>	<i>R</i> ²
Modèle	3	24,10*	< 0,0001	0,29
Nombre de points (N)	1	25,86*	< 0,0001	
Polarité (P)	1	11,68*	0,0008	
N × P	1	34,74*	< 0,0001	
Erreur	176			

* $p < 0,05$

Tableau 4.41 Analyses de contrastes évaluant l'effet du nombre de points de réponse sur les indices de discrimination des items de satisfaction en fonction de la polarité de l'échelle de réponse

Polarité	Nombre de points	Indice de discrimination moyen	$F(1, 176)$	p	ω^2
Bipolaire	6	0,72	0,33	0,5682	0,00
	10	0,71			
Unipolaire	6	0,60	60,28*	< 0,0001	0,24
	10	0,74			

* $p < 0,025$

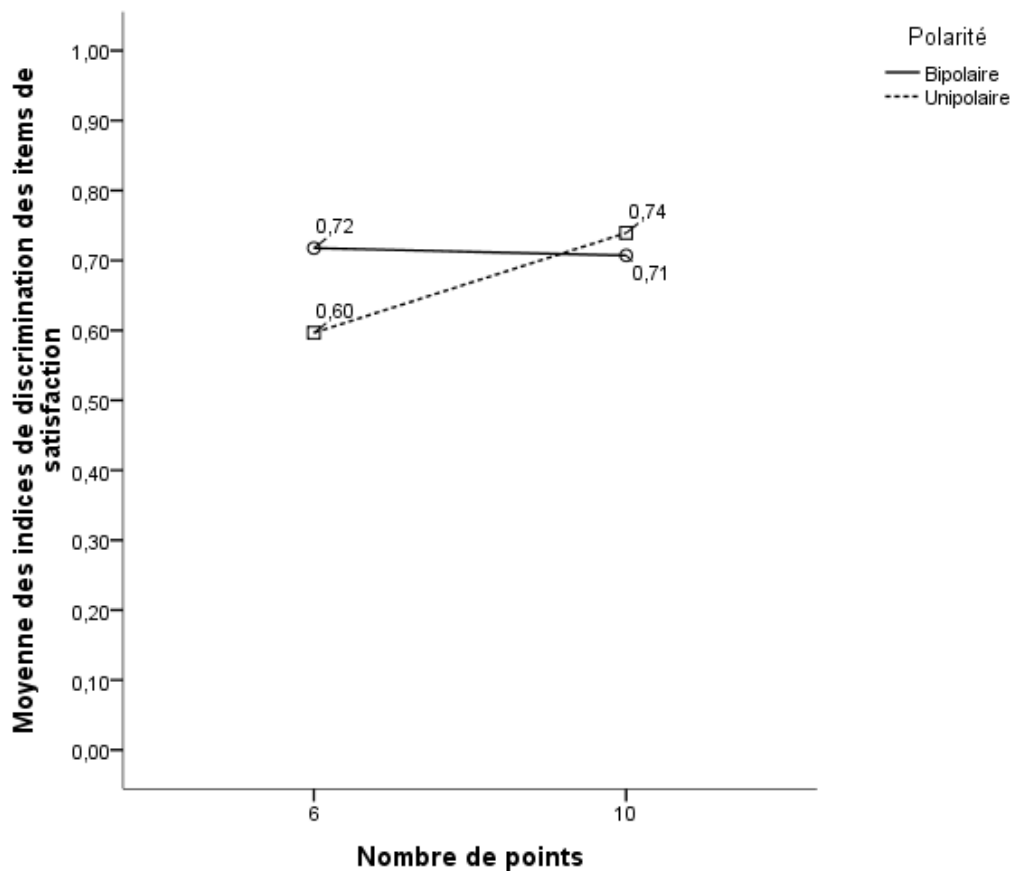


Figure 4.7 Illustration de l'effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse sur les indices de discrimination des items de satisfaction

H_{13} : le coefficient alpha de Cronbach d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points

Le test de Hakstian et Whalen a été utilisé pour vérifier si le coefficient alpha de Cronbach du questionnaire de satisfaction diffère de manière statistiquement significative selon le nombre de points de l'échelle de réponse. Ce test ne permet pas de tester la présence d'un effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse. Cependant, les analyses réalisées pour vérifier H_{12} ont identifié que les effets de ces deux variables sur les indices de discrimination des items de satisfaction interagissent entre eux. En conséquence, nous avons décidé de vérifier l'effet du nombre de points de l'échelle de réponse de manière séparée pour les échelles de réponse bipolaires et unipolaires. Le tableau 4.42 présente les valeurs du coefficient alpha de Cronbach pour le questionnaire de satisfaction en fonction du nombre de points de l'échelle de réponse selon que cette dernière est bipolaire ou unipolaire, d'une part, ainsi que les résultats des tests de Hakstian et Whalen, d'autre part. Ces résultats montrent que le questionnaire de satisfaction dont les items de satisfaction ont une échelle de réponse en dix points plutôt qu'en six points a un coefficient alpha de Cronbach supérieur, mais que ceci s'observe seulement dans la condition des échelles de réponse unipolaires. La treizième hypothèse de recherche ne peut donc être acceptée que pour la condition des échelles de réponse unipolaires.

Tableau 4.42 Coefficients alpha de Cronbach en fonction du nombre de points de l'échelle de réponse selon sa polarité

Polarité	Nombre de points	n	α	Hakstian et Whalen	
				$\chi^2_{(dl=1)}$	p
Bipolaire	6	133	0,9798	0,1215	0,7274
	10	130	0,9785		
Unipolaire	6	130	0,9629	15,7673	0,0001
	10	120	0,9821		

* $p < 0,05$

H_{14} : le degré de validité de convergence d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points

Le test de comparaison basé sur la transformation r à Z de Fisher pour les différences entre deux ou plusieurs coefficients de corrélation indépendants a été utilisé pour tester si le degré de validité de convergence¹³ du questionnaire de satisfaction varie en fonction du nombre de points de l'échelle de réponse. Toutefois, ce test ne permet pas de vérifier s'il existe un effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse. Comme les analyses visant à tester H_{13} ont montré que le nombre de points dans l'échelle de réponse influence la fidélité (*i.e.*, le coefficient alpha de Cronbach) du questionnaire de satisfaction, mais que cet effet dépend de la polarité de l'échelle de réponse, nous avons décidé de tester l'effet du nombre de points de l'échelle de réponse séparément pour les échelles de réponse bipolaires et unipolaires. Le tableau 4.43 présente les indices de validité de convergence du questionnaire de satisfaction en fonction du nombre de points de l'échelle de réponse selon que l'échelle de réponse est bipolaire ou unipolaire, de même que les résultats des tests statistiques réalisés. Ces résultats indiquent que la validité de convergence du questionnaire de satisfaction est supérieure si l'échelle de réponse compte dix points plutôt que six, mais uniquement dans la condition des échelles de réponse unipolaires. Ainsi, la quatorzième hypothèse de recherche est acceptée pour la condition des échelles de réponse unipolaires, mais pas dans celle des échelles de réponse bipolaires.

¹³ L'indice de validité de convergence correspond au coefficient de corrélation entre les mesures multi-items et globale de satisfaction.

Tableau 4.43 Indices de validité de convergence en fonction du nombre de points et de la polarité de l'échelle de réponse

Polarité	Nombre de points	<i>n</i>	<i>r</i> ^a (indice de validité de convergence)	<i>r</i> à <i>z</i> de Fisher	
				$\chi^2_{(df=1)}$	<i>p</i>
Bipolaire	6	154	0,82	0,0748	0,7844
	10	157	0,83		
Unipolaire	6	153	0,75	9,6285*	0,0019
	10	150	0,87		

* $p < 0,05$

^a Ces coefficients de corrélation ont tous une valeur $p < 0,0001$

***H*₁₅ : le degré de validité nomologique d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points**

Les indices de validité nomologique correspondent, rappelons-le, (1) à la covariance entre la satisfaction et l'attitude envers le comportement, (2) au coefficient structurel entre la satisfaction et les intentions comportementales ainsi qu'à (3) l'erreur résiduelle du modèle. Afin de vérifier si la validité nomologique du questionnaire de satisfaction varie selon le nombre de points dans l'échelle de réponse, un test d'invariance a été réalisé pour chacun de ces trois paramètres. La correction de Bonferonni a donc été appliquée pour contrôler la probabilité de commettre une erreur de type 1 de sorte que le seuil de signification statistique a été fixé à 0,017 (0,05/3). De plus, pour tenir compte du fait que les analyses précédentes ont révélé que le nombre de points dans l'échelle de réponse n'a pas le même effet sur la fidélité (*H*₁₃) et la validité de convergence (*H*₁₄) selon que l'échelle de réponse est bipolaire ou unipolaire, nous avons décidé *a priori* de tester *H*₁₅ séparément pour les échelles de réponse bipolaires et unipolaires.

Nous présentons à titre informatif les corrélations entre la satisfaction des clients et les variables associées (*i.e.*, attitude envers la fréquentation du musée, intentions comportementales) en fonction du nombre de points dans l'échelle de réponse du questionnaire de satisfaction, et ce, pour les conditions des échelles de réponse bipolaires et unipolaires (tableau 4.44). Le tableau 4.45 montre les valeurs des trois indices de validité nomologique en fonction du nombre de points de l'échelle de réponse selon que cette dernière est bipolaire ou unipolaire. Les valeurs des indices d'ajustement et du coefficient de détermination correspondants sont présentées à l'annexe 13 (tableau A13.3). Le tableau 4.46, quant à lui, montre les résultats des tests d'invariance pour chacun des trois paramètres structurels en fonction du nombre de points dans l'échelle de réponse pour les conditions des échelles de réponse bipolaires et unipolaires, respectivement. Les résultats de ces tests indiquent qu'aucune des valeurs du chi-carré n'est significative au seuil alpha de 0,017. En d'autres termes, le degré de validité nomologique du questionnaire de satisfaction ne varie pas de manière significative en fonction du nombre de points de l'échelle de réponse. En conséquence, H_{15} ne peut être acceptée.

Tableau 4.44 Corrélations entre la satisfaction des clients et les variables associées en fonction du nombre de points et de la polarité de l'échelle de réponse

Polarité	Nombre de points	n	Attitude	Intention de revenir	Intention de faire du bouche à oreille positif
Bipolaire	6	154	0,42 ^{***}	0,13	0,37 ^{***}
	10	157	0,23 ^{**}	0,25 ^{**}	0,35 ^{***}
Unipolaire	6	153	0,60 ^{***}	0,27 ^{***}	0,55 ^{***}
	10	150	0,32 ^{***}	0,25 ^{**}	0,33 ^{***}

* $p < 0,05$ ** $p < 0,01$ *** $p < 0,001$

Tableau 4.45 Paramètres structurels et coefficient de détermination selon le nombre de points et la polarité de l'échelle de réponse

Polarité	Nombre de points	<i>n</i>	Satisfaction –	Satisfaction –	<i>d</i> ₁
			attitude	intentions comportementales	
Bipolaire	6	154	0,42 (4,33)	0,25 (0,02)	0,89 (1,00)
	10	157	0,23 (2,06)	0,35 (0,02)	0,72 (1,00)
Unipolaire	6	153	0,60 (4,67)	0,40 (0,03)	0,65 (1,00)
	10	150	0,32 (3,17)	0,24 (0,02)	0,78 (1,00)

Note. Les coefficients sont tous significatifs à $p < 0,05$. Les valeurs présentées sont les coefficients standardisés et, entre parenthèses, les coefficients non standardisés.

Tableau 4.46 Tests de l'invariance des trois paramètres structurels selon le nombre de points de réponse pour les échelles de réponse bipolaires et unipolaires

Polarité	Satisfaction - attitude		Satisfaction - intentions comportementales		<i>d</i> ₁	
	χ^2	<i>p</i>	χ^2	<i>p</i>	χ^2	<i>p</i>
	Bipolaire	1,655	0,198	1,550	0,213	3,890
Unipolaire	3,864	0,049	2,430	0,119	1,631	0,202

Mise à l'épreuve des 15 hypothèses de recherche : un résumé

Nous terminons cette présentation des résultats en récapitulant les décisions prises à l'égard des 15 hypothèses de recherche. À cette fin, le tableau 4.47 résume les décisions pour les hypothèses relatives à l'effet de la polarité de l'échelle de réponse (H_1 à H_9). Ce dernier met en évidence que ces neuf premières hypothèses ne sont pas acceptées lorsque le questionnaire de satisfaction a une échelle de réponse en dix points. Par contre, lorsque l'échelle de réponse compte six points, trois hypothèses sont acceptées, mais avec certaines réserves étant donné la présence d'un effet d'interaction entre la polarité et l'étiquetage de l'échelle de réponse. Plus précisément, H_1 et son corolaire, H_4 , sont acceptées, mais uniquement si l'échelle de réponse possède des étiquettes. La troisième hypothèse de recherche (H_3), quant à elle, est acceptée que l'échelle de réponse possède ou pas des étiquettes. Toutefois, par rapport à cette hypothèse, l'interaction entre la polarité et l'étiquetage de l'échelle de réponse indique que la taille d'effet de la polarité sur le degré d'asymétrie des items de satisfaction est plus importante si l'échelle de réponse est étiquetée.

Le tableau 4.48, quant à lui, résume les décisions pour les six hypothèses relatives à l'effet du nombre de points de l'échelle de réponse (H_{10} à H_{15}). L'examen de ce tableau permet de constater qu'aucune de ces six hypothèses de recherche n'est acceptée dans la condition des échelles de réponse bipolaires. En revanche, trois d'entre elles sont acceptées dans la condition des échelles de réponse unipolaires. Plus précisément, dans cette condition, les items de satisfaction ont des indices de discrimination (H_{12}) plus importants si l'échelle de réponse compte dix points plutôt que six. De plus, le questionnaire de satisfaction a un coefficient alpha de Cronbach (H_{13}) et une validité de convergence (H_{14}) supérieurs lorsque l'échelle de réponse compte dix points plutôt que six.

Tableau 4.47 Décisions pour les hypothèses relatives à l'effet de la polarité de l'échelle de réponse (H_1 à H_9)

Hypothèses	Décision		
	6 points		10 points
	Avec étiquettes	Sans étiquettes	
H_1 Les items de satisfaction dont l'échelle de réponse est unipolaire présentent des indices de difficulté inférieurs à ceux des items dont l'échelle de réponse est bipolaire.	Acceptée	Rejetée	Rejetée
H_2 Les items de satisfaction dont l'échelle de réponse est unipolaire présentent des écarts-types supérieurs aux items dont l'échelle de réponse est bipolaire.	Rejetée	Rejetée	Rejetée
H_3 Les items de satisfaction dont l'échelle de réponse est unipolaire présentent des indices d'asymétrie moins importants que les items dont l'échelle de réponse est bipolaire.	Acceptée	Acceptée	Rejetée
H_4 La satisfaction exprimée dans un questionnaire dont l'échelle de réponse est unipolaire est moins élevée que celle exprimée dans un questionnaire dont l'échelle de réponse est bipolaire.	Acceptée	Rejetée	Rejetée
H_5 La variance de la satisfaction exprimée dans un questionnaire dont l'échelle de réponse est unipolaire est plus grande que celle exprimée dans un questionnaire dont l'échelle de réponse est bipolaire.	Rejetée	Rejetée	Rejetée
H_6 Les indices de discrimination des items de satisfaction dont l'échelle de réponse est unipolaire sont supérieurs à ceux dont l'échelle de réponse est bipolaire	Rejetée	Rejetée	Rejetée
H_7 Le coefficient alpha de Cronbach d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celui d'un questionnaire dont l'échelle de réponse est bipolaire.	Rejetée	Rejetée	Rejetée
H_8 Le degré de validité de convergence d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celui d'un questionnaire dont l'échelle de réponse est bipolaire.	Rejetée	Rejetée	Rejetée
H_9 Le degré de validité nomologique d'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire est supérieur à celle d'un questionnaire dont l'échelle de réponse est bipolaire.	Rejetée	Rejetée	Rejetée

Tableau 4.48 Décisions pour les hypothèses relatives à l'effet du nombre de points dans l'échelle de réponse (H_{10} à H_{15})

Hypothèses		Décision	
		Bipolaires	Unipolaires
H_{10}	Les items de satisfaction dont l'échelle de réponse a dix points présentent des coefficients de variation supérieurs aux items dont l'échelle de réponse a six points.	Rejetée	Rejetée
H_{11}	La variance de la satisfaction exprimée dans un questionnaire dont l'échelle de réponses a dix points est plus élevée que celle obtenue dans un questionnaire dont l'échelle de réponse a six points.	Rejetée	Rejetée
H_{12}	Les indices de discrimination des items de satisfaction dont l'échelle de réponse a dix points sont supérieurs à ceux dont l'échelle de réponse a six points.	Rejetée	Acceptée
H_{13}	Le coefficient alpha de Cronbach d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points.	Rejetée	Acceptée
H_{14}	Le degré de validité de convergence d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points.	Rejetée	Acceptée
H_{15}	Le degré de validité nomologique d'un questionnaire de satisfaction dont l'échelle de réponse a dix points est supérieur à celui d'un questionnaire dont l'échelle de réponse a six points.	Rejetée	Rejetée

Le chapitre qui suit discute de ces résultats et de leurs implications. Il dégage les principaux constats liés à notre objectif de recherche et les examine à la lumière de la littérature pertinente.

Chapitre 5 : Discussion

L'objectif de cette étude était d'éprouver l'effet de la polarité (bipolaire c. unipolaire), du nombre de points (six c. dix) et, à titre exploratoire, de l'étiquetage (*i.e.*, avec ou sans étiquettes) d'une échelle de réponse sur la distribution des réponses à des items de satisfaction et sur les qualités psychométriques d'un questionnaire de satisfaction. À cette fin, nous avons émis 15 hypothèses de recherche que nous avons testées au chapitre précédent à l'aide de deux blocs d'analyses statistiques. Le premier bloc d'analyses visait à vérifier les neuf hypothèses concernant la polarité de l'échelle de réponse (H_1 à H_9), alors que le second servait à tester les six hypothèses relatives au nombre de points dans l'échelle de réponse (H_{10} à H_{15}). Certains des résultats obtenus sont allés dans le sens de nos hypothèses de recherche, alors que d'autres n'ont pas permis de les soutenir. Dans ce chapitre, nous discutons d'abord des résultats relatifs à l'effet de la polarité de l'échelle de réponse et ensuite de ceux quant à l'effet du nombre de points de l'échelle de réponse. Nous poursuivons avec un examen des constats et implications des résultats de l'étude. Nous terminons cette discussion en abordant les limites de l'étude.

Effet de la polarité de l'échelle de réponse (H_1 à H_9)

Rappelons d'entrée de jeu que les hypothèses relatives à la polarité de l'échelle de réponse, H_1 à H_9 , ont été testées séparément pour les échelles de réponse en six et en dix points. Pour les échelles de réponse en six points, les résultats diffèrent selon que l'échelle de réponse comporte des étiquettes ou pas. Ainsi, trois des neuf hypothèses ont été acceptées pour les échelles de réponse avec étiquettes (H_1 , H_3 et H_4) et une seule dans le cas des échelles de réponse sans étiquettes (H_3). Ci-après, nous discutons de l'effet de la polarité de l'échelle de réponse selon que cette dernière est étiquetée ou pas, respectivement.

Tout d'abord, lorsque l'échelle de réponse comporte des étiquettes, les résultats vont dans le sens de trois de nos hypothèses de recherche, c'est-à-dire qu'ils démontrent qu'utiliser une échelle de réponse unipolaire plutôt que bipolaire fait baisser la moyenne des items de satisfaction (H_1) et celle du score global au questionnaire de satisfaction (H_4) et réduit le degré d'asymétrie négative des items de satisfaction (H_3). Ces résultats concordent avec l'idée voulant que plus l'échelle de réponse fait référence à la partie positive du continuum évalué, ce continuum étant dans ce cas-ci la satisfaction des clients, plus l'item sera difficile à endosser et moins les réponses seront concentrées à l'extrémité positive de l'échelle de réponse (French-Lazovik & Gibson, 1984; Gannon & Ostrom, 1996; Lam & Klockars, 1982; Vocational Psychology Research, 2007; Weiss *et al.*, 1967). Ainsi, il apparaît que cette proposition théorique s'applique bien aux items de satisfaction et que les répondants trouvent ces items plus difficiles à endosser lorsque l'échelle de réponse (avec étiquettes) est unipolaire plutôt que bipolaire, raison pour laquelle ils ont moins tendance à choisir les réponses les plus favorables.

Par contre, sans en avoir fait la démonstration, plusieurs chercheurs ont aussi rapporté que ces items plus difficiles à endosser devraient avoir des écarts-types plus importants (Vocational Psychology Research, 2007; Weiss *et al.*, 1967) et un plus grand pouvoir discriminant (French-Lazovik & Gibson, 1984; Gannon & Ostrom, 1996; Lam & Klockars, 1982) lorsqu'employés avec un groupe d'individus se situant majoritairement dans la partie positive du continuum psychologique. Nous avons donc émis les hypothèses voulant que les items de satisfaction dont l'échelle de réponse est unipolaire plutôt que bipolaire ont des écarts-types (H_2) et des indices de discrimination (H_6) dont les valeurs sont plus grandes. En corollaire, nous avons également proposé qu'un questionnaire de satisfaction dont l'échelle de réponse est unipolaire plutôt que bipolaire a une variance plus grande (H_5) de même que des indices fidélité (H_7) et de validité (H_8 et H_9) supérieurs. Toutefois, les résultats ne supportent aucune de ces hypothèses et n'appuient donc pas nos hypothèses. En effet, même si les items de satisfaction (avec étiquettes) sont plus difficiles à endosser et ont une asymétrie négative moins importante lorsque l'échelle de réponse est unipolaire plutôt que bipolaire, cela n'affecte aucunement leurs écarts-types et leurs qualités

psychométriques. Le fait que nous n'ayons pas observé de relation entre, d'une part, les indices de difficulté des items de satisfaction et, d'autre part, l'ampleur de leurs écarts-types et de leurs indices de discrimination peut sans doute s'expliquer par l'influence plutôt faible de la polarité de l'échelle de réponse sur les indices de difficulté des items de satisfaction. En effet, les items de satisfaction dont l'échelle de réponse (avec étiquettes) est unipolaire plutôt que bipolaire ne sont que légèrement plus difficiles à endosser, leur indice moyen de difficulté n'étant inférieur que de seulement 6,45 points de pourcentage (73,04 c. 79,49), ce qui ne représente que 0,32 point de moins sur l'échelle de réponse en six points (4,65 c. 4,97). Il apparaît plausible de croire qu'une si petite différence soit sans effet sur les écarts-types et les indices de discrimination des items de satisfaction. Qui plus est, les items de satisfaction demeurent faciles à endosser et les scores globaux de satisfaction restent élevés même lorsque l'échelle de réponse est unipolaire.

Au moins deux hypothèses peuvent être avancées pour expliquer que la polarité de l'échelle de réponse (avec étiquettes) a peu d'effet sur la difficulté des items de satisfaction, celles-ci n'étant pas mutuellement exclusives. La première hypothèse repose sur les idées de Wildt et Mazis (1978) ainsi que de Klockars et Yamagishi (1988). Ces chercheurs avancent que pour interpréter une étiquette apposée sur un point dans une échelle de réponse, les répondants utilisent deux informations : ladite étiquette et l'endroit où se retrouve cette étiquette dans l'échelle de réponse (*i.e.*, proche de l'extrémité négative, du centre ou de l'extrémité positive de l'échelle de réponse). Selon leur théorie, une même étiquette, par exemple « plutôt satisfait », serait perçue plus négativement et serait plus facile à endosser si elle est placée loin de l'extrémité positive de l'échelle de réponse et, inversement, serait perçue plus positivement et serait plus difficile à endosser si elle est placée proche de l'extrémité positive de l'échelle de réponse. Si on applique cette idée aux résultats observés pour les items de satisfaction, on peut penser que même si les étiquettes de l'échelle de réponse unipolaire réfèrent à la partie positive du continuum de satisfaction, les étiquettes situées au centre et à l'extrémité négative de l'échelle de réponse seraient perçues plus négativement et seraient plus faciles à endosser qu'on pourrait le penser d'emblée. Pour

cette raison, employer une échelle de réponse unipolaire plutôt que bipolaire aurait peu d'effet sur les indices de difficulté des items de satisfaction.

Une deuxième hypothèse peut aussi être avancée pour expliquer que la polarité de l'échelle de réponse (avec étiquettes) a peu d'effet sur les indices de difficulté des items de satisfaction. Elle consiste en ce que les répondants pourraient interpréter les premiers points de l'échelle de réponse unipolaire (aucunement satisfait, très peu satisfait, passablement satisfait) pas uniquement comme des degrés nuls à faibles de satisfaction, mais aussi comme des degrés plus ou moins importants d'insatisfaction. Par exemple, on peut penser qu'une personne qui se dit très peu ou passablement satisfaite exprime également par cette réponse un certain degré d'insatisfaction. Les résultats des études de Davern et Cummins (2006) et de Mazaheri et Theuns (2009) vont également dans cette direction. Ceux-ci ont demandé à des participants d'inscrire leur degré de satisfaction sur une échelle de réponse unipolaire représentant la satisfaction, par exemple allant de « Aucunement satisfait » à « Extrêmement satisfait », et d'inscrire leur degré d'insatisfaction sur une échelle de réponse unipolaire représentant l'insatisfaction, par exemple allant de « Aucunement insatisfait » à « Extrêmement insatisfait ». Dans ces deux études, les degrés d'insatisfaction et de satisfaction étaient corrélés de manière importante ($-0,92 \geq r \geq -0,67$), c'est-à-dire que plus les scores sur l'échelle de satisfaction étaient faibles, plus les scores sur l'échelle d'insatisfaction étaient élevés et, inversement, plus les scores sur l'échelle de satisfaction étaient élevés, plus les scores sur l'échelle d'insatisfaction étaient faibles. De surcroît, dans l'étude de Davern et Cummins, la somme des moyennes des scores de satisfaction et d'insatisfaction était approximativement de 100 %. Par exemple, si pour un élément évalué le score moyen de satisfaction était de 75 %, alors le score moyen d'insatisfaction se situait à environ 25 %. À partir de ces résultats, ces chercheurs ont conclu que la satisfaction et l'insatisfaction sont simplement la réciproque l'une de l'autre et que si l'une décroît l'autre croît automatiquement. A la lumière de ces résultats, les échelles de réponse de satisfaction qualifiées d'unipolaires ne le seraient pas véritablement, car elles aussi mesureraient l'insatisfaction. Elles n'offriraient ainsi pas beaucoup plus de choix au répondant pour préciser son degré de satisfaction par rapport à une échelle de réponse bipolaire. Si cette

hypothèse s'avérerait juste, elle impliquerait qu'il n'existe pas de point neutre sur le continuum de satisfaction qui sépare la satisfaction de l'insatisfaction et où l'individu possède à la fois un degré nul de satisfaction et un degré nul d'insatisfaction, souvent exprimé par l'étiquette « ni satisfait ni insatisfait » dans les échelles de réponse bipolaires comprenant un point milieu. Le centre du continuum de satisfaction serait plutôt un amalgame d'insatisfaction et de satisfaction et le passage de l'un à l'autre se ferait graduellement.

Par ailleurs, dans la condition où l'échelle de réponse n'est pas étiquetée, les résultats démontrent que le fait d'utiliser une échelle unipolaire plutôt que bipolaire diminue effectivement le degré d'asymétrie négative des items de satisfaction, comme le prévoyait H_3 . Par contre, dans cette condition, aucun autre résultat ne va dans le sens de nos hypothèses de recherche concernant la polarité de l'échelle de réponse. Ces résultats sont quelque peu déstabilisants à première vue, car on pourrait s'attendre à ce que des items de satisfaction ayant un degré d'asymétrie moins important soient aussi plus difficiles à endosser, comme le prévoyait H_1 , et aient des écarts-types plus grands, comme le prédisait H_2 . De surcroît, non seulement les items de satisfaction dont l'échelle de réponse (sans étiquettes) est unipolaire ont des indices de difficulté similaires à ceux dont l'échelle de réponse est bipolaire, mais en plus leurs écarts-types sont inférieurs à ceux dont l'échelle de réponse est bipolaire, ce qui va à l'inverse de H_2 . De même, leurs indices de discrimination sont inférieurs, tout comme le degré de fidélité du questionnaire, ceci, inversement à ce que prévoient respectivement H_6 et H_7 . Une interprétation possible de ces résultats est que, en l'absence d'étiquettes, l'échelle de réponse unipolaire serait plus difficile à interpréter que l'échelle de réponse bipolaire. En effet, lorsque l'échelle de réponse est étiquetée, le fait d'employer une échelle de réponse unipolaire au lieu d'une qui est bipolaire n'affecte ni les écarts-types ni les qualités psychométriques des items et du questionnaire de satisfaction. Ainsi, il se pourrait que les répondants soient incertains de la signification des points de l'échelle de réponse unipolaire lorsqu'il n'y a pas d'étiquettes sur chacun de ces points. Il leur serait alors plus difficile d'y situer leur degré de satisfaction avec exactitude et avec consistance pour les différents items du questionnaire de satisfaction. Cette incertitude

pourrait être attribuable au fait que les répondants sont moins familiers avec les échelles de réponse unipolaires que bipolaires, ces dernières étant plus populaires. Elle pourrait aussi être due au fait qu'une échelle de réponse (sans étiquettes) unipolaire serait plus difficile à décoder qu'une échelle de réponse (sans étiquettes) bipolaire étant donné qu'elle n'est pas symétrique. En effet, contrairement à l'échelle de réponse bipolaire, ses deux moitiés ne représentent pas simplement les moitiés positive et négative du continuum de satisfaction. Il devient alors plus difficile pour le répondant de comprendre à quel degré de satisfaction correspond chacun des points de l'échelle de réponse. Qui plus est, cette difficulté pourrait être encore plus grande pour les premiers points de l'échelle de réponse unipolaire (sans étiquettes) étant donné qu'ils représentent à la fois la moitié négative de l'échelle de réponse et la partie positive du continuum de satisfaction. Les résultats semblent d'ailleurs soutenir cette hypothèse voulant que la moitié négative de l'échelle de réponse unipolaire (sans étiquettes) ait une signification ambiguë pour les répondants. Effectivement, les items de satisfaction dont l'échelle de réponse est unipolaire (sans étiquettes) ont des degrés d'asymétrie négative moins importants et des écarts-types inférieurs à ceux dont l'échelle de réponse est bipolaire (sans étiquettes), bien que leurs indices de difficulté (moyennes) soient comparables. Autrement dit, les répondants assignés à la condition où les échelles de réponse étaient unipolaires sans étiquettes ont choisi moins fréquemment les premiers points de l'échelle de réponse que ceux assignés à la condition où les échelles de réponse étaient bipolaires sans étiquettes.

En ce qui a trait aux échelles de réponse en dix points, aucune des neuf hypothèses relatives à l'effet de la polarité n'a pu être acceptée dans cette condition. Le fait qu'il n'y ait aucune différence associée à la polarité de l'échelle de réponse aux plans des paramètres des items de satisfaction (indice de difficulté (H_1), écart-type (H_2), indice d'asymétrie (H_3), indice de discrimination (H_6)) et des caractéristiques psychométriques du questionnaire de satisfaction (moyenne (H_4), variance (H_5), fidélité (H_7), validité (H_8 et H_9)) permet de penser que les gens répondent de manière similaire que l'échelle de réponse soit unipolaire ou bipolaire lorsque cette dernière comprend dix points. Autrement dit, lorsque l'échelle de réponse compte dix points, sa polarité n'affecterait pas le comportement de réponse des

individus. Nous suggérons deux explications à ces résultats. La première est que les répondants porteraient davantage attention aux chiffres (1 jusqu'à 10) qu'aux étiquettes apposées aux extrémités de l'échelle de réponse dans le cas d'une échelle de réponse en dix points. La seconde est que les répondants se seraient sentis plus familiers avec l'échelle de réponse en dix points qu'avec celle en six points, ce pourquoi ils auraient observé moins attentivement l'échelle de réponse en dix points. Ils n'auraient alors pas remarqué si cette dernière est bipolaire ou unipolaire, utilisant les étiquettes uniquement pour identifier les minimum et maximum de l'échelle de réponse. En effet, la notation sur dix points est plus fréquemment utilisée que la notation sur six pour évaluer les biens et services (p. ex. dans la presse écrite, dans les magazines télévisés et dans les sites Internet destinés aux consommateurs et visant à les aider à faire des choix de consommation éclairés), d'une part, et les études en psychologie ont démontré que les individus observent moins attentivement les stimuli familiers (p. ex. les visages de personnes qu'ils connaissent bien, les objets qu'ils utilisent dans leur quotidien) que les stimuli nouveaux (p. ex. l'apparence d'une personne qu'on leur présente pour la première fois, un nouvel immeuble dans leur quartier, un lieu qu'ils fréquentent peu souvent) (Solomon *et al.*, 1999), d'autre part.

Lorsque comparés aux résultats d'autres études portant sur d'autres concepts que la satisfaction et utilisant une méthodologie différente, les présents résultats permettent de soulever l'hypothèse que l'effet de la polarité de l'échelle de réponse, en plus de différer selon l'étiquetage et le nombre de points de l'échelle de réponse, pourrait aussi différer selon le mode d'administration du questionnaire ou selon le concept mesuré par l'item. En effet, en utilisant une échelle de réponse allant de 0 à 10, Davern et Cummins (2006) ainsi que Mazaheri et Theuns (2009) ont trouvé que la polarité de l'échelle de réponse n'affecte pas la difficulté d'un item, tout comme la présente étude, alors que O'Muircheartaigh *et al.* (1995) ont trouvé qu'utiliser une échelle de réponse unipolaire plutôt que bipolaire augmente la difficulté d'un item. Dans les deux premières études, les chercheurs ont utilisé des items de satisfaction et un questionnaire auto-administré, tout comme nous. Par contre, dans la seconde étude, les chercheurs ont utilisé un item mesurant l'opinion des individus sur un aspect d'ordre politique, soit le pouvoir accordé à un organisme gouvernemental

régissant la publicité, et ce, à l'aide d'un questionnaire téléphonique. Ainsi, du moins dans le cas d'une échelle de réponse allant de 0 à 10 ou de 1 à 10, la polarité d'une telle échelle de réponse pourrait avoir une influence plus ou moins importante selon que les étiquettes apposées aux extrémités réfèrent au concept de satisfaction ou à un autre concept. Il se pourrait également que les répondants accordent plus d'importance aux étiquettes apposées aux extrémités de l'échelle de réponse, et conséquemment à sa polarité, si le questionnaire est administré par téléphone. En effet, en l'absence d'un questionnaire pour leur servir de support visuel, les répondants n'ont d'autre choix que de porter attention et de mémoriser les informations transmises par l'intervieweur concernant l'échelle de réponse, par exemple les étiquettes apposées à ses extrémités.

Effet du nombre de points de l'échelle de réponse (H_{10} à H_{15})

Avant de discuter de l'effet du nombre de points de l'échelle de réponse (six c. dix), rappelons que les hypothèses s'y rattachant (H_{10} à H_{15}) ont été testées uniquement dans la condition des échelles de réponse sans étiquettes. Pour tester ces hypothèses, nous avons tenu compte de la polarité de l'échelle de réponse dans nos analyses étant donné que les résultats concernant l'effet de la polarité laissaient soupçonner la présence d'un effet d'interaction entre le nombre de points et la polarité de l'échelle de réponse. Les résultats ont d'ailleurs confirmé la présence de cet effet d'interaction. Ceux-ci vont en partie dans le sens de nos hypothèses dans la condition où les échelles de réponse unipolaires sont utilisées, mais pas dans la condition où on utilise les échelles de réponse bipolaires. Dans le cas des échelles de réponse unipolaires (sans étiquettes), trois de nos six hypothèses ont été acceptées. Plus précisément, l'augmentation du nombre de points de réponse est associée à une augmentation des indices de discrimination des items de satisfaction (H_{12}) ainsi que de la fidélité (H_{13}) et de la validité de convergence (H_{14}) du questionnaire de satisfaction. Par contre, dans la condition des échelles de réponse bipolaires (sans étiquettes), aucune de nos hypothèses n'a été acceptée, c'est-à-dire que le fait d'employer une échelle de réponse en dix points plutôt qu'en six points n'est pas associé à des coefficients de variation (H_{10}) et

des indices de discrimination (H_{12}) supérieurs des items de satisfaction et est sans effet sur les caractéristiques du questionnaire de satisfaction : variance (H_{11}), fidélité (H_{13}), validité (H_{14} et H_{15}).

Nous pourrions conclure de cette interaction entre le nombre de points et la polarité de l'échelle de réponse que le fait d'augmenter le nombre de points de six à dix permet aux répondants d'exprimer avec plus de justesse leur degré de satisfaction, mais ce uniquement lorsque l'échelle de réponse est unipolaire. Cette conclusion nous apparaît cependant difficile à soutenir, car il serait ardu d'expliquer pourquoi c'est le cas pour les échelles de réponses unipolaires, mais pas pour celles bipolaires. Une conclusion plus plausible est que nous observons cette interaction non pas parce que l'effet du nombre de points de réponse sur les qualités psychométriques des items et du questionnaire de satisfaction est modéré par celui de la polarité de l'échelle de réponse, mais parce que l'effet de la polarité est modéré par celui du nombre de points de l'échelle de réponse. En effet, nos résultats montrent que la polarité de l'échelle de réponse affecte les qualités psychométriques des items et du questionnaire de satisfaction ayant des échelles de réponse en six points (sans étiquettes), mais pas lorsque les échelles de réponse comportent dix points. Plus spécifiquement, dans la condition où les échelles de réponse en six points sont utilisées, le recours à une échelle de réponse unipolaire plutôt que bipolaire est associé à des qualités psychométriques inférieures des items et du questionnaire de satisfaction. Comme nous l'avons soulevé précédemment, une explication possible à ces résultats est que l'échelle de réponse en six points et sans étiquettes pourrait être plus difficile à interpréter lorsque celle-ci est unipolaire plutôt que bipolaire, d'une part, et que, d'autre part, les répondants pourraient ne pas remarquer la polarité de l'échelle de réponse lorsque cette dernière compte dix points. En d'autres termes, l'amélioration des qualités psychométriques observée dans la condition des échelles de réponse unipolaires ne serait pas attribuable au fait qu'un plus grand nombre de points permet aux individus d'être plus précis dans leurs réponses, mais au fait qu'augmenter de six à dix le nombre de points d'une échelle de réponse unipolaire (sans étiquettes) ferait simplement disparaître la difficulté d'interpréter une échelle de réponse unipolaire sans étiquettes. Cette hypothèse nous apparaît plus juste,

car elle explique pourquoi, dans la condition des échelles de réponse bipolaires, les qualités psychométriques du questionnaire de satisfaction et de ses items sont semblables que l'échelle de réponse compte six ou dix points.

Si cette hypothèse s'avère exacte, elle impliquerait que, indépendamment de la polarité de l'échelle de réponse, les répondants ne précisent pas de manière plus fine leur degré de satisfaction lorsque l'échelle de réponse compte dix points plutôt que six, contrairement à ce que certains auteurs avaient suggéré. D'ailleurs, en ce sens, leurs réponses aux items de satisfaction ont une variabilité égale ou inférieure lorsqu'il y a plus de points dans l'échelle de réponse, comme en témoignent les coefficients de variations (H_{10}). Ces observations pourraient s'expliquer de plusieurs façons. Par exemple, les individus pourraient trouver difficile de différencier entre eux chacun des dix points de l'échelle de réponse ou simplement en être incapables, car l'exercice de différenciation pour les répondants se complexifie à mesure que le nombre de choix de réponse augmente (Cox, 1980). Par ailleurs, Alwin (1997) mentionne que les répondants à un questionnaire pourraient être capables de différencier entre eux les dix points d'une échelle de réponse, mais ne pas être motivés à fournir l'effort cognitif nécessaire pour y arriver. En ce sens, il a été avancé que plusieurs répondants ne cherchent pas à répondre à un questionnaire de manière optimale, c'est-à-dire au meilleur de leur capacité, par exemple en étant attentifs à chacune des questions, en tentant de comprendre chacun des choix de réponse et en essayant d'identifier le meilleur d'entre eux, soit celui qui représente le plus exactement possible la réponse à ce qu'on leur demande. Plusieurs répondants utiliseraient plutôt une stratégie cognitive visant à alléger ce travail et se contenteraient de répondre de manière adéquate à un questionnaire, c'est-à-dire qu'ils fourniraient un effort minimalement suffisant pour comprendre l'essentiel de la question et des choix de réponse afin de donner la réponse qu'ils estiment adéquate, soit celle qui représente suffisamment bien la réponse à ce qu'on leur demande, sans que ce soit nécessairement celle correspondant le plus à ce qu'ils pensent (Alwin & Krosnick, 1991; Krosnick, 1991; Krosnick & Alwin, 1987). Autrement dit, plusieurs répondants feraient un compromis entre la qualité de leur réponse et l'effort à fournir pour répondre. On peut d'ailleurs émettre l'hypothèse que cette stratégie est plus fréquente dans

le cadre d'un questionnaire à faible implication pour la personne, c'est-à-dire où la qualité de ses réponses est sans conséquence pour elle, comme c'est le cas pour un questionnaire de satisfaction, que dans celui d'un questionnaire à haute implication, tel qu'un questionnaire d'entrevue de sélection en vue d'un emploi.

Un dernier point doit être abordé avant de terminer cette discussion concernant l'effet du nombre de points composant une échelle de réponse. Rappelons que nous avons relevé trois études ayant évalué l'effet du nombre de points dans l'échelle de réponse sur un aspect des qualités psychométriques d'un questionnaire de satisfaction. Parmi elles, deux ont rapporté des qualités psychométriques supérieures en utilisant une échelle de réponse en dix (échelle de 1 à 10) ou onze (échelle de 0 à 10) points plutôt qu'en cinq ou sept points (Alwin, 1997; Coelho & Esteves, 2007), ce qui va à l'encontre des résultats de la présente étude. Ces résultats divergents pourraient être dus au fait que dans ces deux études, l'échelle de réponse servant de point de comparaison comprenait un nombre impair de points de réponse, c'est-à-dire cinq ou sept avec un point milieu, alors que la nôtre comprenait un nombre pair de choix de réponse. De surcroît, ces chercheurs ne précisent pas si les différences qu'ils rapportent en faveur de l'échelle de réponse en dix/onze points sont statistiquement significatives. Elles pourraient donc n'être que le fruit du hasard, d'autant plus que la troisième étude réalisée à l'aide d'un questionnaire de satisfaction n'a trouvé aucune amélioration des qualités psychométriques liée à l'utilisation d'une échelle de réponse en dix points (Neumann & Neumann, 1981). De plus, les études qui n'ont pas utilisé une mesure de la satisfaction arrivent généralement à la conclusion que peu ou pas de gains peuvent être réalisés en utilisant plus de sept points de réponse.

Constats et implications

Nous présentons ici les principaux constats qui peuvent être faits par rapport à l'objectif de cette étude. Ce faisant, nous identifions comment ces nouvelles connaissances s'inscrivent dans les trois champs de recherche touchés par cette thèse, soit la mesure de la satisfaction des clients, les aspects cognitifs des méthodes de sondage et la psychométrie. Enfin, nous dégagons les implications théoriques et pratiques qui découlent de l'étude.

En premier lieu, plusieurs constats contribuent à l'avancement des connaissances relativement à la mesure de la satisfaction des clients. En effet, c'est la première fois qu'une étude évalue de manière aussi complète les suggestions faites par certains auteurs d'utiliser dans un questionnaire de satisfaction une échelle de réponse unipolaire plutôt que bipolaire ou une échelle de réponse en dix points au lieu d'une en comprenant cinq, six ou sept.

Soulignons tout d'abord les constats qui se dégagent relativement à la proposition d'utiliser une échelle de réponse unipolaire plutôt que bipolaire pour mesurer la satisfaction des clients. Nos résultats indiquent qu'utiliser une échelle de réponse unipolaire plutôt que bipolaire ne permet pas d'améliorer les qualités psychométriques d'un questionnaire de satisfaction et que cela pourrait même les diminuer légèrement, et ce, contrairement à ce que certains chercheurs avaient pensé. L'échelle de réponse bipolaire, celle utilisée le plus couramment dans les enquêtes et sondages, apparaît donc être le meilleur choix pour mesurer la satisfaction des clients étant donné qu'elle permet d'obtenir des qualités psychométriques égales ou supérieures à celles obtenues en employant une échelle de réponse unipolaire. Il faut tout de même nuancer cette conclusion en rappelant que les qualités psychométriques d'un questionnaire de satisfaction et de ses items demeurent très bonnes peu importe que l'échelle de réponse soit bipolaire ou unipolaire. En ce sens, le

recours à une échelle de réponse unipolaire ne semble pas entraîner de conséquences négatives importantes.

D'autres constats se dégagent par rapport à la suggestion qu'ont faite certains chercheurs d'utiliser une échelle de réponse en dix points plutôt qu'en cinq, six ou sept points afin de permettre aux répondants d'avoir plus de latitude pour préciser leur degré de satisfaction et, conséquemment, d'augmenter le pouvoir discriminant de l'instrument de mesure. En effet, rappelons que bien qu'il soit généralement admis que le nombre idéal de points dans une échelle de réponse de type Likert se situe entre cinq et sept, certains chercheurs ont émis l'hypothèse que ce nombre pourrait être de dix (échelle de 1 à 10) ou onze (échelle de 0 à 10) points dans le cas d'un questionnaire de satisfaction étant donné la fréquence élevée de scores élevés. Cependant, contrairement à l'hypothèse de ces chercheurs, il semble d'après nos résultats que les questionnaires de satisfaction, à l'instar des questionnaires évaluant d'autres concepts psychologiques, ne bénéficient pas au plan psychométrique d'une telle augmentation du nombre de points de l'échelle de réponse. En d'autres termes, augmenter de six à dix le nombre de points dans l'échelle de réponse ne semble pas améliorer le pouvoir discriminant des items de satisfaction ni la fidélité et la validité du questionnaire de satisfaction. Ces résultats signifient qu'utiliser une échelle de réponse en six points est suffisant et représente un choix parcimonieux. Néanmoins, utiliser une échelle de réponse en dix points ne nuit pas aux qualités psychométriques des items et d'un questionnaire de satisfaction. De plus, selon l'étude de Preston et Colman (2000), les répondants préféreraient une échelle de réponse en dix points aux échelles de réponse comprenant moins de dix points pour évaluer les différents éléments de la qualité de service d'une entreprise. Pour cette raison, l'échelle de réponse en dix points pourrait tout de même s'avérer un choix judicieux, car elle pourrait permettre d'obtenir un meilleur taux de questionnaires dûment complétés ou encore un meilleur taux de réponse, un avantage tant pour le chercheur que pour l'entreprise.

La présente étude permet donc de constater que le nombre de points, la polarité et l'étiquetage de l'échelle de réponse ne sont pas des caractéristiques qui influencent de manière déterminante les qualités psychométriques d'un questionnaire de satisfaction. Conséquemment, le chercheur ou l'entreprise ne devrait pas porter une trop grande attention à ces trois caractéristiques lors de la construction d'un questionnaire de satisfaction.

Toujours concernant la mesure de la satisfaction des clients, l'étude nous apprend que l'asymétrie négative observée dans les réponses à des items de satisfaction ne semble pas nuire de manière importante aux qualités psychométriques d'un questionnaire de satisfaction et de ses items. D'une part, elle ne semble pas avoir nui à la fidélité du questionnaire de satisfaction et au pouvoir discriminant de ses items. Effectivement, le coefficient alpha de Cronbach du questionnaire de satisfaction et les indices de discrimination de ses items sont élevés malgré le fait que les coefficients d'asymétrie des items de satisfaction aient des valeurs moyennes se situant, selon l'échelle de réponse, entre -0,65 et -1,32. Ces résultats empiriques viennent appuyer ceux obtenus à l'aide de simulations Monte Carlo réalisées par Greer, Dunlap, Hunter et Bernan (2006). Ces chercheurs avaient trouvé que l'asymétrie des items d'un instrument de mesure affecte à la baisse les estimations des corrélations inter-items et du coefficient alpha de Cronbach surtout lorsque cette asymétrie est importante (*i.e.*, coefficients d'asymétrie égaux ou supérieurs à 2,0), qu'elle va en sens opposé selon les items (*i.e.*, certains items ont une asymétrie négative alors que d'autres items présentent une asymétrie positive), que le nombre d'items est petit et que leur corrélation inter-items moyenne est faible. Par contre, dans un questionnaire de satisfaction, l'asymétrie des scores est toujours négative et, dans la présente étude, 96 % des coefficients d'asymétrie estimés pour les items de satisfaction étaient inférieurs à 2,0. Ces résultats permettent de penser que la concentration de réponses positives dans les questionnaires de satisfaction ne représenterait pas une menace sérieuse à leur fidélité et au pouvoir discriminant de leurs items si leur formulation est claire et leur contenu pertinent. D'autre part, elle ne semble pas non plus affecter de manière importante les indices de validité d'un questionnaire de satisfaction, puisque les validités de

convergence et nomologique ont pu être démontrées même si les scores de satisfaction avaient une asymétrie négative. Cette conclusion par rapport à l'effet de l'asymétrie des scores sur la validité d'un questionnaire de satisfaction doit cependant être nuancée, car nous ne savons pas dans quelle mesure les valeurs des indices de validité observés auraient été similaires ou supérieures si les scores de satisfaction avaient eu une distribution symétrique. Néanmoins, il demeure que les résultats de l'étude renforcent l'idée qu'un questionnaire de satisfaction aura de bonnes qualités psychométriques si son ou ses concepteurs respectent les règles de l'art en psychométrie, c'est-à-dire les étapes généralement recommandées pour construire un instrument de mesure possédant de bons degrés de validité et de fidélité.

La principale menace associée à l'asymétrie négative des scores de satisfaction pèserait donc principalement sur l'estimation des relations entre la satisfaction des clients et les variables ayant une distribution asymétrique positive. En effet, c'est lorsque les asymétries des scores vont en sens opposés que la corrélation entre deux variables est la plus sous-estimée (Greer *et al.*, 2006). Par exemple, il est reconnu que les personnes ayant un revenu élevé fréquentent davantage les attractions culturelles comme les musées que celles ayant un faible revenu (Harrison & Shaw, 2004; Rogers, 1998). Ainsi, on pourrait s'attendre à ce qu'il y ait une association entre la satisfaction envers les musées et le revenu des individus, une variable possédant une distribution asymétrique positive. Ces deux variables possédant des distributions allant en sens opposés, il y a lieu de croire que la force de leur association serait sous-estimée.

Le fait que l'asymétrie négative des scores de satisfaction ne semble pas nuire de manière importante aux qualités psychométriques d'un questionnaire de satisfaction a aussi des implications pour les recherches futures portant sur la mesure de la satisfaction des clients. Plutôt que de tenter d'identifier des méthodes qui favoriseraient l'augmentation du pouvoir discriminant des items, comme cela fut le cas dans cette thèse, les travaux devraient se concentrer sur les méthodes à privilégier pour atténuer les problèmes d'estimation

statistique que peut entraîner la présence d'asymétrie dans les scores de satisfaction. Entre autres choses, il serait intéressant d'examiner dans quelle mesure les méthodes statistiques dites plus robustes à la non-normalité permettent d'obtenir de meilleures estimations des relations entre la satisfaction et d'autres variables. À cet égard, certains auteurs ont suggéré d'estimer les paramètres des modèles de régression ou structurels à l'aide de la méthode des moindres carrés partiels (*Partial Least Squares*) (Fornell, 1992) ou du maximum de vraisemblance robuste (*Robust Maximum Likelihood*) (O'Loughlin & Coenders, 2004). De même, comme les transformations mathématiques visant à normaliser la distribution des variables rendent plus difficile l'interprétation des résultats, des méthodes alternatives à ces transformations pourraient être explorées. À titre d'exemple, la modélisation des réponses aux items de satisfaction à l'aide des modèles de type Rasch, qui possèdent un paramètre de normalisation des scores des individus, diminuerait sans doute l'ampleur de l'asymétrie des scores de satisfaction. Des méthodes très simples pourraient aussi être envisagées, telles que de regrouper *a posteriori* les réponses d'insatisfaction en un seul choix de réponse, celles-ci étant de toute façon peu fréquentes. Par exemple, dans une échelle de réponse bipolaire en six points allant de 1 « Extrêmement insatisfait » à 6 « Extrêmement satisfait », les points les moins utilisés par les répondants pourraient être regroupés de manière à diminuer l'asymétrie négative et ainsi faciliter les analyses statistiques.

La présente recherche permet aussi d'approfondir les connaissances acquises dans les études portant sur les aspects cognitifs des méthodes de sondage. Contrairement à la méthodologie employée dans ces études, la présente recherche a testé plusieurs combinaisons de polarité, de nombre de points et d'étiquetage de l'échelle de réponse. L'étude des combinaisons de ces trois facteurs a permis d'apprendre que l'effet de la polarité de l'échelle de réponse sur le comportement de réponse aux items est difficile à prédire. Effectivement, il n'est pas le même selon que l'échelle de réponse est étiquetée ou pas et diffère en fonction du nombre de points de l'échelle de réponse. De plus, comme nous l'avons mentionné plus tôt dans cette discussion, l'effet de la polarité de l'échelle de réponse pourrait aussi dépendre du concept mesuré ou du mode d'administration du questionnaire. Ces résultats permettent d'expliquer pourquoi certaines études ont trouvé

qu'utiliser une échelle de réponse unipolaire plutôt que bipolaire augmente la difficulté des items, alors que d'autres n'ont pas observé cet effet. Effectivement, nous avons trouvé que la polarité de l'échelle de réponse affecte les indices de difficulté des items uniquement si l'échelle de réponse est étiquetée. Nous comprenons maintenant mieux pourquoi les chercheurs qui ont utilisé une échelle de réponse étiquetée ont observé qu'utiliser une échelle de réponse unipolaire plutôt que bipolaire augmente la difficulté des items (*i.e.*, Farmer *et al.*, 2001; French-Lazovik & Gibson, 1984; Vocational Psychology Research, 2007), alors que trois des quatre études qui ont employé une échelle de réponse sans étiquettes n'ont pas observé cette relation (*i.e.*, Davern & Cummins, 2006; Follman, 1976; Mazaheri & Theuns, 2009).

En dernier lieu, notons que cette recherche contribue également aux connaissances en psychométrie. Effectivement, à notre connaissance, c'est la première fois qu'une étude s'intéresse à l'effet de la polarité de l'échelle de réponse sur les qualités psychométriques d'un instrument de mesure. Jusqu'à ce jour, ce sont surtout les effets du nombre de points et de l'étiquetage de l'échelle de réponse qui ont été étudiés. En ce sens, l'étude nous apprend que la polarité de l'échelle de réponse est une variable qui semble affecter très peu les qualités psychométriques d'un instrument de mesure.

Limites de l'étude

Finalement, avant de terminer cette discussion et de conclure cette thèse, il importe d'aborder les limites de l'étude. Les premières limites concernent la validité interne des résultats. Tout d'abord, nous avons utilisé un devis quasi-expérimental et il est possible avec ce type de devis qu'un ou plusieurs des groupes de comparaison créés aient des caractéristiques différentes des autres. Lorsque c'est le cas, ces différences, plutôt que les variables indépendantes étudiées, peuvent être à l'origine des effets observés dans une recherche (Kirk, 1995). Ensuite, notre taux de réponse de 57 % constitue un second facteur

nuisant à la validité interne des résultats. En effet, nous ne savons pas si le pourcentage des individus (43 %) qui ont refusé de participer à l'étude diffère significativement de ceux qui ont accepté. Si nous avions eu un plus grand taux de réponse, les résultats auraient peut-être été différents. Notamment, les variances des items et des scores de satisfaction auraient pu être différents et leur estimation plus stable avec une plus grande taille d'échantillon. De même, les estimés des qualités psychométriques de la mesure de satisfaction et de ses items auraient été plus précis si le nombre de participants dans chacune des six conditions du plan factoriel incomplet avait été plus important, par exemple si un ratio de dix participants par item de satisfaction avait été utilisé. De la sorte, les résultats auraient été encore plus fiables. Une telle exigence aurait cependant été difficile à respecter. Pour reprendre le même exemple, un ratio de dix participants par item de satisfaction aurait nécessité 450 répondants par groupe de comparaison, ce qui représente une taille totale d'échantillon de 2 700 participants. Pour atteindre cette taille d'échantillon, il aurait été nécessaire de solliciter près de 10 000 visiteurs du Musée de la civilisation. De plus, selon Osborne (2008), environ 40 % des études en psychologie utilisent un ratio inférieur ou égal à cinq participants par item, comme dans la présente étude.

Une autre limite à la validité interne des résultats est que les variables indépendantes « nombre de points de l'échelle de réponse » et « étiquetage de l'échelle de réponse » n'ont pu être croisées à cause de la difficulté d'élaborer des étiquettes pour l'ensemble des points des échelles de réponse en dix points. Pour cette raison, nous avons dû adopter un plan factoriel incomplet, ce qui a eu pour conséquence de nous empêcher de tester simultanément dans un même modèle statistique les trois effets étudiés (la polarité, le nombre de points et l'étiquetage de l'échelle de réponse). Comme seulement une ou deux de ces trois variables étaient entrées à la fois dans nos modèles statistiques, les estimations de leurs paramètres ne tiennent pas compte de l'effet de la ou des variables absentes du modèle. De même, il n'a pas été possible de tester l'effet d'interaction entre l'étiquetage et le nombre de points de l'échelle de réponse ainsi que la triple interaction entre les trois variables d'intérêt (polarité, nombre de points et étiquetage de l'échelle de réponse). Il est donc impossible de savoir si l'étiquetage de l'échelle de réponse a un effet dans le cas des

échelles de réponse en dix points ou encore s'il existe une triple interaction entre la polarité, le nombre de points et l'étiquetage de l'échelle de réponse. Toutefois, comme les échelles de réponse en dix points avec des étiquettes sont très rarement utilisées, ces résultats auraient été moins pertinents.

Les autres limites concernent la validité externe des résultats. Dans un premier temps, les résultats observés dans la condition des échelles de réponse avec des étiquettes sont en partie tributaires des étiquettes que nous avons choisies pour identifier les différents points des échelles de réponse bipolaire et unipolaire. En ce sens, nos résultats concernant les échelles de réponse étiquetées ne peuvent être généralisés à des échelles de réponse qui utiliseraient des adverbes différents sur leurs points de réponse. À titre d'exemple, nos échelles de réponse bipolaire et unipolaire avec étiquettes avaient toutes deux « Extrêmement satisfait » comme maximum. Cependant, il arrive que les échelles de réponse bipolaires se terminent par un adverbe moins fort, généralement « Très satisfait ». Lorsque c'est le cas, il est probable que l'écart entre les indices de difficulté des items de satisfaction dont l'échelle de réponse est unipolaire plutôt que bipolaire soit plus important que celui observé ici. Dans un second temps, notre échantillon de participants était composé d'individus relativement scolarisés par rapport à la moyenne de la population, une caractéristique typique des visiteurs du Musée de la civilisation (Daignault, 2008). En effet, parmi ceux qui étaient aux études, la majorité (52 %) poursuivait des études de niveau universitaire, et parmi ceux qui avaient terminé leurs études, plus des deux tiers (68 %) possédaient un diplôme universitaire. Comme il est possible que le comportement de réponse à un questionnaire soit influencé par le niveau de scolarité d'un individu, il se pourrait que les résultats de l'étude ne soient pas généralisables à des populations moins scolarisées.

En somme, différents facteurs peuvent avoir influencé la justesse des résultats et leur généralisabilité. Ces derniers doivent donc être interprétés en tenant compte de ces limites. Néanmoins, le fait que certains des résultats de cette étude rejoignent ceux rapportés dans la

littérature scientifique permet de penser que ces limites n'affectent pas outre mesure les principaux constats.

Conclusion

Généralement, dans les études évaluant le degré de satisfaction des clients, les réponses d'insatisfaction sont peu fréquentes et les scores de satisfaction présentent une moyenne élevée, ce pourquoi leur distribution est caractérisée par une asymétrie négative. Plusieurs auteurs pensent que cette distribution asymétrique typique des scores de satisfaction entraîne une sous-estimation de la variance et de la covariance des outils utilisés pour mesurer ce concept et, en conséquence, que ceci a un impact négatif sur leurs qualités psychométriques. Dans ce contexte, certains auteurs ont proposé d'augmenter la sensibilité des échelles de réponse des mesures de satisfaction afin d'augmenter la variance associée à la caractéristique mesurée (*i.e.*, la variance vraie), soit la satisfaction. La première suggestion à cette fin est d'utiliser une échelle de réponse unipolaire, qui se concentre uniquement sur la partie positive du continuum de satisfaction, plutôt que bipolaire, qui couvre autant la partie négative (l'insatisfaction) que la partie positive (la satisfaction) de ce continuum. La seconde proposition est d'opter pour une échelle de réponse en dix ou onze points, c'est-à-dire allant de 1 à 10 ou de 0 à 10, plutôt que pour une échelle de réponse plus traditionnelle en cinq, six ou sept points.

Découlant de cette problématique, l'objectif de cette étude était d'évaluer l'effet de la polarité (bipolaire c. unipolaire), du nombre de points (six c. dix) et, à titre exploratoire, de l'étiquetage (*i.e.*, avec ou sans étiquettes) de l'échelle de réponse sur la distribution des réponses et sur les qualités psychométriques d'un questionnaire de satisfaction des clients et de ses items. Au plan théorique, l'étude s'est appuyée sur l'approche cognitive des méthodes de sondage, un courant d'études s'intéressant au comportement de réponse à un questionnaire. Par ailleurs, au plan méthodologique, cette recherche a employé un devis quasi-expérimental, plus précisément un plan factoriel incomplet à six conditions, et un échantillon de 907 participants.

À l'égard de la polarité de l'échelle de réponse, l'étude nous apprend que recourir à une échelle de réponse unipolaire plutôt que bipolaire peut, selon que l'échelle de réponse possède ou pas des étiquettes, respectivement, n'avoir aucun effet ou avoir un effet négatif sur les écarts-types et les indices de discrimination des items de satisfaction ainsi que sur les qualités psychométriques d'un questionnaire de satisfaction des clients. Nos résultats ne soutiennent donc pas l'idée selon laquelle une échelle de réponse unipolaire plutôt que bipolaire permet de mieux discriminer entre eux les sujets quant à leur degré de satisfaction et ainsi améliorer les qualités psychométriques d'un questionnaire de satisfaction et de ses items. Ainsi, comme le choix d'une échelle de réponse unipolaire plutôt que bipolaire peut avoir un effet négatif sur les qualités psychométriques d'un questionnaire de satisfaction, nous recommandons d'utiliser une échelle de réponse bipolaire pour ce type de questionnaire. Par ailleurs, les résultats montrent aussi que l'effet de la polarité de l'échelle de réponse est difficile à prédire, car il diffère selon que l'échelle de réponse possède ou pas des étiquettes et selon qu'elle compte six ou dix points.

En ce qui a trait à l'effet du nombre de points de l'échelle de réponse, nous concluons des résultats de l'étude qu'utiliser une échelle de réponse en dix points plutôt qu'en six n'est pas associé à une plus grande variabilité des réponses aux items de satisfaction et des scores globaux de satisfaction, d'une part, et est sans effet sur les qualités psychométriques d'un questionnaire de satisfaction des clients et de ses items, d'autre part. En ce sens, les répondants ne précisent pas de manière plus fine leur degré de satisfaction lorsque l'échelle de réponse compte dix points plutôt que six, contrairement à ce que certains auteurs avaient suggéré. Dans ce contexte, l'échelle de réponse en six points représente un choix parcimonieux, mais il demeure qu'utiliser une échelle de réponse en dix points ne nuit pas aux qualités psychométriques des items et d'un questionnaire de satisfaction.

De manière plus générale, nos résultats suggèrent que les trois caractéristiques d'une échelle de réponse étudiées ici (*i.e.*, la polarité, le nombre de points et l'étiquetage) ont une influence marginale sur les qualités psychométriques d'un questionnaire de satisfaction, ce

qui rejoint certains résultats rapportés dans la littérature. Conséquemment, ces trois caractéristiques de l'échelle de réponse devraient avoir une importance secondaire dans le processus de construction d'un questionnaire de satisfaction des clients. L'étude permet également de constater que les qualités psychométriques d'un questionnaire de satisfaction et de ses items, incluant leur pouvoir de discrimination, sont tout à fait acceptables malgré la distribution asymétrique négative des scores de satisfaction. En ce sens, l'étude suggère que la distribution asymétrique négative des scores de satisfaction ne semble pas représenter une menace importante pour les qualités psychométriques d'un questionnaire de satisfaction.

Nous terminons avec quelques suggestions pour les prochaines études. D'abord, il serait intéressant d'examiner les effets de la polarité et du nombre de points de l'échelle de réponse d'un questionnaire de satisfaction des clients sur ses qualités psychométriques à l'aide d'un autre cadre d'analyse que celui de la théorie classique des tests pour voir dans quelle mesure les résultats convergent avec ceux observés ici. Les modèles de la théorie des réponses aux items, comme le *Rating Scale Model* (Andrich, 1978), pourraient notamment être employés. Ensuite, en plus de s'intéresser à l'échelle de réponse, il serait pertinent de vérifier comment la manière d'opérationnaliser le concept de satisfaction peut affecter la distribution des scores et les qualités psychométriques d'une mesure de satisfaction des clients. Un autre facteur qui pourrait affecter la distribution des réponses obtenues à un questionnaire de satisfaction et les qualités psychométriques d'un tel questionnaire est la formulation des items, notamment leur intensité sémantique. Ce facteur serait particulièrement pertinent à étudier lorsque la satisfaction des clients est mesurée à l'aide d'une échelle de réponse où le répondant doit indiquer son degré d'accord ou de désaccord. Enfin, il serait utile d'explorer les méthodes, statistiques ou non, qui pourraient être utilisées pour améliorer la justesse des estimations statistiques malgré la distribution asymétrique négative des scores de satisfaction.

Bibliographie

- ACSI. (2009). The American Customer Satisfaction Index. Accédé le 9 décembre, 2009, de <http://www.theacsi.org/>
- Ajzen, I. (2002). Constructing a theory of planned behavior questionnaire: conceptual and methodological considerations. Accédé le 1er février 2008, de <http://people.umass.edu/aizen/pdf/tpb.measurement.pdf>
- Ajzen, I., Brown, T. C., & Carvajal, F. (2004). Explaining the discrepancy between intentions and actions: the case of hypothetical bias in contingent valuation. *Personality and Social Psychology Bulletin*, 30(9), 1108-1121.
- Ajzen, I., Brown, T. C., & Rosenthal, L. H. (1996). Information bias in contingent valuation: effects of personal relevance, quality of information, and motivational orientation. *Journal of Environmental Economics and Management*, 30, 43-57.
- Alain, M. (2002). *Réussir la performance des services aux clients*. Montréal: Éditions nouvelles AMS.
- Allaire, A. (2000). *Résultats généraux de l'enquête de mars 2000 auprès des visiteurs adultes du Musée de la civilisation*. Québec, QC: Musée de la civilisation de Québec.
- Alwin, D. F. (1992). Information transmission in the survey interview: number of response categories and the reliability of attitude measurement. *Sociological Methodology*, 22, 83-118.
- Alwin, D. F. (1997). Feeling thermometers versus 7-point scales: which are better? *Sociological Methods & Research*, 25(3), 318-340.
- Alwin, D. F., & Krosnick, J. A. (1991). The reliability of survey attitude measurement: the influence of question and respondent attributes. *Sociological Methods & Research*, 20(1), 139-181.
- Amoo, T., & Friedman, H. H. (2001). Do numeric scale values influence responses to rating scales? *Journal of International Marketing and Marketing Research*, 26, 41-46.
- Anastasi, A. (1988). *Psychological testing* (6e éd.). New York, NY: Macmillan.
- Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7e éd.). Upper Saddle River, NJ: Prentice Hall.
- Anderson, E. W., Fornell, C., & Lehmann, D. R. (1994). Customer satisfaction, market share, and profitability: findings from Sweden. *Journal of Marketing*, 58(3), 53-66.

- Anderson, E. W., & Sullivan, M. W. (1993). The antecedents and consequences of customer satisfaction for firms. *Marketing Science*, 12(2), 125-143.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561-573.
- Applebaum, R., Straker, J. K., & Geron, S. (2000). *Assessing satisfaction in health and long-term care : practical approaches to hearing the voices of consumers*. New York: Springer.
- Armstrong, R. D., Jones, D. H., & Wang, Z. (1998). Optimization of classical reliability in test construction. *Journal of Educational and Behavioral Statistics*, 23(1), 1-17.
- Athanassopoulos, A. D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of Business Research*, 47(3), 191-207.
- Athiyaman, A. (1997). Linking student satisfaction and service quality perceptions: the case of university education. *European Journal of Marketing*, 31(7), 528-540.
- Baker, D. A., & Crompton, J. L. (2000). Quality, satisfaction, and behavioral intentions. *Annals of Tourism Research*, 27(3), 785-804.
- Bandalos, D. L., & Enders, C. K. (1996). The effects of nonnormality and number of response categories on reliability. *Applied Measurement in Education*, 9(2), 151 - 160.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Bartikowski, B., Chandon, J. L., & Gierl, H. (2005). *L'utilisation des échelles de mesure à l'international: une étude interculturelle*. France: Institut d'Administration des Entreprises.
- Bearden, W. O., & Teel, J. E. (1983). Selected determinants of consumer satisfaction and complaint reports. *Journal of Marketing Research*, 20(1), 21-28.
- Berg, J. E., Bergeron, J., Seitz, J. A., Monroe, M. C., & McConnell, L. (2007, November 14-17). *Understanding reliability when measuring environmental knowledge in elementary students*. Document présenté à: North American Association for Environmental Education (NAEE) Research Symposium, Virginia Beach, VA.
- Birkett, N. J. (1986). *Selecting the number of response categories for likert-type scale*. Document présenté à: Proceedings of the Survey Research Methods Section, American Statistical Association.

- Blais, J.-G., & Grondin, J. (2008). *A study of the influence of labels associated with anchor points of Likert-type response scales in survey questionnaires*. Document présenté à: International Objective Measurement Workshop (IOMW).
- Bloemer, J. M. M., & Kasper, H. D. P. (1995). The complex relationship between consumer satisfaction and brand loyalty. *Journal of Economic Psychology*, 16, 311-329.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model : fundamental measurement in the human sciences* (2e éd.). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Boone, W. J., & Britt, R. S. (2001). Museum visitor's attitudes toward exhibits, staffing, and amenities. Dans R. Adams (Éd.), *Museum visitor services manual* (pp. 67-71). Washington, DC: American Association of Museums.
- Bruner, G. I., & Hensel, P. (1993). Multi-item scale usage in marketing journals: 1980 to 1989. *Journal of the Academy of Marketing Science*, 21(4), 339-344.
- Bryant, W. K. (1990). *The economic organization of the household*. Cambridge: Cambridge University Press.
- Busacca, B., & Padula, G. (2005). Understanding the relationship between attribute performance and overall satisfaction: theory, measurement and implications. *Marketing Intelligence & Planning*, 23(6), 543-561.
- Cardozo, R. N. (1965). An experimental study of customer effort, expectation, and satisfaction. *Journal of Marketing Research*, 2(August), 244-249.
- CEFRIQ. (2003). *NetGouv: Sondage réalisé auprès des citoyens, des entreprises et des travailleurs autonomes du Québec*. Québec, QC: CEFRIQ.
- Chakrapani, C. (1998). *How to measure service quality & customer satisfaction : the informal field-guide for tools and techniques*. Chicago, Ill.: American Marketing Association.
- Christian, L. M., & Dillman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly*, 68(1), 57-80.
- Chu, R. (2002). Stated-importance versus derived-importance customer satisfaction measurement. *The Journal of Services Marketing*, 16(4), 285-301.
- Churchill, G. A., Jr., & Peter, J. P. (1984). Research design effects on the reliability of rating scales: a meta-analysis. *Journal of Marketing Research*, 21(4), 360-375.
- Churchill, G. A., & Surprenant, C. (1982). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research*, 19(November), 491-504.

- Churchland, P. M. (1992). *A neurocomputational perspective: the nature of mind and the structure of science*. Cambridge, MA: MIT Press.
- Coelho, P. S., & Esteves, S. P. (2007). The choice between a five-point and a ten-point scale in the framework of customer satisfaction measurement. *International Journal of Market Research*, 49(3), 313-339.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7(3), 249-253.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2e éd.). Hillsdale, NJ: L. Erlbaum Associates.
- Cooper, A. R., Cooper, M. B., & Duban, D. F. (1989). Measurement instrument development using two competing concepts of customer satisfaction. *Journal of consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 2, 28-35.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104.
- Council of Better Business Bureaus. (2001). Effective customer relations dependable tips and helpful information. Accédé le 04/10, 2006, de <http://www.bbb.org/alerts/article.asp?ID=449>
- Courville, T. G. (2004). *An empirical comparison of item response theory and classical test theory item/person statistics*. Thèse de doctorat, Texas A&M University, College Station.
- Cox, E. P. I. (1980). The optimal number of response alternatives for a scale: a review. *Journal of Marketing Research*, 17(4), 407-422.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Cronbach, L. J. (1969). *Essentials of psychological testing* (3e éd.). New York: Harper & Row.
- Cronin, J. J., Brady, M. K., & Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing*, 76(2), 193-218.
- Cummins, R. A. (1995). On the trail of the gold standard for subjective well-being. *Social Indicators Research*, 35(2), 179-200.
- Cummins, R. A. (2002). *Vale ComQol: caveats to using the Comprehensive Quality of Life Scale*.

- Cummins, R. A. (2003). Normative life satisfaction: measurement issues and a homeostatic model. *Social Indicators Research, 64*(2), 225.
- Cummins, R. A., & Gullone, E. (2000). *Why we should not use 5-point Likert scales: the case for subjective quality of life measurement*. Document présenté à: Second International Conference on Quality of Life in Cities, Singapore: National University of Singapore.
- Czepiel, J. A., & Rosenberg, L. J. (1977). Consumer satisfaction: concept and measurement. *Journal of the Academy of Marketing Science, 5*(4), 403-411.
- Daignault, L. (2008). *L'enquête générale de public au Musée de la civilisation - Été 2008*. Québec: Service de la recherche et de l'évaluation - Musée de la civilisation.
- Danaher, P. J. (1997). Using conjoint analysis to determine the relative importance of service attributes measured in customer satisfaction surveys. *Journal of Retailing, 73*(2), 235-260.
- Danaher, P. J., & Haddrell, V. (1996). A comparison of question scales used for measuring customer satisfaction. *International Journal of Service Industry Management, 7*(4), 4-26.
- Davern, M., & Cummins, R. A. (2006). Is life dissatisfaction the opposite of life satisfaction? *Australian Journal of Psychology, 58*(1), 1-7.
- Day, R. L. (1984). Modeling choices among alternative responses to dissatisfaction. *Advances in Consumer Research, 11*, 496-499.
- DeVellis, R. F. (1991). *Scale development: theory and applications*. Thousand Oaks, CA: SAGE Publications.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffon, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71-75.
- Dixon, P. N., Bobo, M., & Stevick, R. A. (1984). Response differences and preferences for all-category-defined and end-defined Likert formats. *Educational and Psychological Measurement, 44*(1), 61-66.
- Dunlap, W. P., Burke, M. J., & Greer, T. (1995). The effect of skew on the magnitude of product-moment correlations. *Journal of General Psychology, 122*(4), 365-377.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.
- Engel, J., Kollat, D., & Blackwell, R. (1968). *Consumer behavior*. New York: Holt, Rinehart and Winston.

- England, B. L. (2003). *Relationship of motivations, decisions making, and satisfaction in museum visitor behavior*. University of Florida.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavioral Research Methods, Instruments, & Computers*, 28, 1-11.
- Farmer, W. L., Thompson, R. C., Heil, S. K. R., & Heil, M. C. (2001). *Latent trait theory analysis of changes in item response anchors*. Washington, D.C.: Office of Aviation Medicine.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: an introduction to theory and research*. Reading, MA: Addison-Wesley Publishing Co.
- Follman, J. (1976). Modes and formats in student rating scales of teaching effectiveness *Applied Psychology*, 25(2), 115-119.
- Fornell, C. (1992). A national satisfaction barometer: the Swedish experience. *Journal of Marketing*, 56(1), 1-21.
- Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American Customer Satisfaction Index: nature, purpose, and findings. *Journal of Marketing*, 60(4), 7-18.
- Fowler, F. J. (1995). *Improving survey questions: design and evaluation* (Vol. 38). Thousand Oaks, CA: SAGE Publications.
- Fowler, F. J. (2002). *Survey research methods* (3e éd. Vol. 1). Thousand Oaks, CA: SAGE Publications.
- Frank, R., Salzman, K., & Fergus, E. (1977). Correlates of consumer satisfaction with outpatient therapy assessed by postcards. *Community Mental Health Journal*, 13, 37-45.
- French-Lazovik, G., & Gibson, C. L. (1984). Effects of verbally labeled anchor points on the distributional parameters of rating measures. *Applied Psychological Measurement* 8(1), 49-57.
- Freud, S., Strachey, J., & Richards, A. (1974). *Introductory lectures on psychoanalysis*. Harmondsworth: Penguin.
- Frisbie, D. A., & Brandenburg, D. C. (1979). Equivalence of questionnaire items with varying response formats. *Journal of Educational Measurement*, 16(1), 43-48.
- Ganglmair, A., & Lawson, R. (2003). Measuring affective response to consumption using Rasch modeling. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 16, 198.

- Gannon, K. M., & Ostrom, T. M. (1996). How meaning is given to rating scales: the effects of response language on category activation. *Journal of Experimental Social Psychology, 32*, 337-360.
- Giese, J. L., & Cote, J. (2000). Defining consumer satisfaction. *Academy of Marketing Science Review, 2000*, 1-27.
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3e éd.). Boston: Allyn and Bacon.
- Gould, G. (1995). Why it is customer loyalty that counts (and how to measure it). *Managing Service Quality, 5*(1), 15-19.
- Greer, T., Dunlap, W. P., Hunter, S. T., & Berman, M. E. (2006). Skew and internal consistency. *Journal of Applied Psychology, 91*(6), 1351-1358.
- Groupe SECOR. (1996). *Démarche d'amélioration de la qualité de service - Volet 3 : rapport d'enquêtes*: Musée d'art contemporain de Montréal et Musée de la civilisation (Québec).
- Guilford, J. P. (1954). *Psychometric methods* (2e éd.). New York, NY: McGraw-Hill.
- Ha, C. L. (1998). The theory of reasoned action applied to brand loyalty. *The Journal of Product and Brand Management, 7*(1), 51.
- Hakstian, A. R., & Whalen, T. E. (1976). A k-sample significance test for independent alpha coefficients. *Psychometrika, 41*(2), 219-231.
- Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International Journal of Service Industry Management, 7*(4), 27.
- Halstead, D., Hartman, D., & Schmidt, S. L. (1994). Multisource effect on the satisfaction formation process. *Journal of the Academy of Marketing Science, 22*(2), 114-129.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: principles and applications*. Norwell, MA: Kluwer Academic Publishers.
- Harrison, P., & Shaw, R. (2004). Consumer satisfaction and post-purchase intentions: an exploratory study of museum visitors. *International Journal of Arts Management, 6*(2), 23-32.
- Hayes, B. E. (1998). *Measuring customer satisfaction: survey design, use, and statistical analysis methods* (2e éd.). Milwaukee, Wis.: ASQC Quality Press.
- Hazelrigg, L. E., & Hard, M. A. (2000). Scaling the semantics of satisfaction. *Social Indicators Research, 49*(2), 147.

- Hennig-Thurau, T., & Klee, A. (1997). The impact of customer satisfaction and relationship quality on customer retention: a critical reassessment and model development. *Psychology & Marketing, 14*(8), 737–764.
- Hill, N., & Alexander, J. (2000). *Handbook of Customer Satisfaction and Loyalty Measurement*: Gower Publishing.
- Hippler, H.-J., & Schwarz, N. (1987). Response effects in surveys. Dans H.-J. Hippler, N. Schwarz & S. Sudman (Éds.), *Social information processing and survey methodology* (pp. 102-123). New York, NY: Springer-Verlag.
- Hooper-Greenhill, E. (2001). A welcome in a new place. Dans R. Adams (Éd.), *Museum visitor services manual* (pp. 16-17). Washington, DC: American Association of Museums.
- Horovitz, J. (2000). Measurement for improved performance. Dans *The seven secrets of service strategy* (pp. 41-56). Harlow, England: Prentice Hall.
- Howard, J. A., & Sheth, J. N. (1969). *The theory of buyer behavior*. New York: Wiley.
- Hrubes, D., Ajzen, I., & Daigle, J. (2001). Predicting hunting intentions and behavior: an application of the theory of planned behavior. *Leisure Sciences, 23*, 165-178.
- Hunt, H. K. (1977). *Conceptualization and measurement of consumer satisfaction and dissatisfaction*. Cambridge, MA: Marketing Science Institute.
- Huo, Y., & Miller, D. (2007). Satisfaction measurement of small tourism sector (museum): Samoa. *Asia Pacific Journal of Tourism Research, 12*(2), 103-117.
- Ironson, G. H., Smith, P. C., Brannick, M. T., Gibson, W. M., & Paul, K. B. (1989). Construction of a Job in General scale: a comparison of global, composite, and specific measures. *Journal of Applied Psychology, 74*(2), 193-200.
- Jobe, J. B., Tourangeau, R., & Smith, A. F. (1993). Contributions of survey research to the understanding of memory. *Applied Cognitive Psychology, 7*(7), 567-584.
- Johns, N., & Clark, S. L. (2001). Customer perception auditing: a means of monitoring the service provided by museums and galleries. Dans R. Adams (Éd.), *Museum visitor services manual* (pp. 17-22). Washington, DC: American Association of Museums.
- Keegan, W. J., & Seringhaus, F. H. R. (1999). *Global marketing management* (2e éd.). Scarborough, Ont.: Prentice-Hall Canada.
- Kirk, R. E. (1995). *Experimental design : procedures for the behavioral sciences* (3e éd.). Pacific Grove, CA: Brooks/Cole.
- Klockars, A. J., & Yamagishi, M. (1988). The influence of labels and positions in rating scales. *Journal of Educational Measurement 25*(2), 85-96.

- Komorita, S. S., & Graham, W. K. (1965). Number of scale points and the reliability of scales. *Educational and Psychological Measurement*, 25(4), 987-995.
- Korn, R. (2001). What visitors think about visitor services. Dans R. Adams (Éd.), *Museum visitor services manual* (pp. 65-66). Washington, DC: American Association of Museums.
- Koskas, J.-C. (1997). Satisfaction - Fidélité. *Revue Française du Marketing*, 164, 89-101.
- Kotler, P., Armstrong, G., & Cunningham, P. H. (1999). *Principles of marketing* (4e éd.). Scarborough, Ont.: Prentice-Hall Canada.
- Kowalski, C. J. (1972). On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. *Applied Statistics*, 21, 1-12.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213-236.
- Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, 51(2), 201.
- LaBarbera, P. A., & Mazursky, D. (1983). A longitudinal assessment of consumer satisfaction/dissatisfaction: the dynamic aspect of cognitive standardization. *Journal of Marketing Research*, 20(4), 394-404.
- Lam, T. C. M., & Klockars, A. J. (1982). Anchor point effects on the equivalence of questionnaire items. *Journal of Educational Measurement*, 19(4), 317-322.
- Larsen, D. L., Attkinson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: development of a general scale. *Evaluation and Program Planning*, 2(3), 197-207.
- Likert, R. (1932). A technique for measurement of attitudes. *Archives of Psychology*, 140(June), 1-55.
- Linn, L. S. (1975). Factors Associated with Patient Evaluation of Health Care. *The Milbank Memorial Fund Quarterly. Health and Society*, 53(4), 531-548.
- Lissitz, & Green. (1975). Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach. *Journal of Applied Psychology*, 60(1), 10-13.
- Mackoy, R. D., & Spreng, R. A. (1995). The dimensionality of consumer satisfaction/dissatisfaction: an empirical examination. *Journal of consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 8, 53-58.
- Maddox, N. R. (1981). Two-factor theory and consumer satisfaction: replication and extension. *Journal of Consumer Research*, 8(June), 97-102.

- Mano, H., & Oliver, R. L. (1993). Assessing the dimensionality and structure of the consumption experience: evaluation, feeling, and satisfaction. *Journal of Consumer Research*, 20(3), 451-466.
- Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: reliability and validity. *Journal of Applied Psychology*, 31(3), 657-674.
- Mazaheri, M., & Theuns, P. (2005). *Bias in overall life satisfaction and life dissatisfaction ratings due to the type of response formats*. Document présenté à: IX Social and Health Sciences Methodology Conference.
- Mazaheri, M., & Theuns, P. (2009). Effects of varying response formats on self-ratings of life-satisfaction. *Social Indicators Research*, 90(3), 381-395.
- McKinney, V., Yoon, K., & Zahedi, F. M. (2002). The measurement of web-customer satisfaction: an expectation and disconfirmation approach. *Information Systems Research*, 13(3), 296-315.
- Medin, D., Ross, B. H., & Markman, A. B. (2005). Memory. Dans *Cognitive psychology* (4e éd.). Hoboken, NJ: John Wiley & Sons.
- Meehan, C. (2002). Measuring satisfaction meaningfully. *Insite*, April/May, 6-7.
- Mellor, D., Cummins, R. A., & Loquet, C. (1999). The gold standard for life satisfaction: confirmation and elaboration using an imaginary scale and qualitative interview. *International Journal of Social Research Methodology*, 2(4), 263-278.
- Musée du Louvre. (2006). *La fréquentation du musée du Louvre: chiffres clés 2006*. Paris: Direction des Publics: Service Études, Évaluation et Prospective.
- Neill, G., Tremblay, M.-È., Végiard, S., Lavoie, A., & Moisan, K. (2007). *Enquête sur la satisfaction des usagers à l'égard des services de santé et des services sociaux du Québec, 2006-2007 : description et méthodologie.*: Institut de la statistique du Québec.
- Neumann, L., & Neumann, Y. (1981). Comparison of six lengths of rating scales: students' attitudes toward instruction. *Psychological Reports*, 48, 399-404.
- Nguyen, T. D., Attkisson, C. C., & Stegner, B. L. (1983). Assessment of patient satisfaction: development and refinement of a Service Evaluation Questionnaire. *Evaluation and Program Planning*, 6(3-4), 299-313.
- Nowacki, M. M. (2005). Evaluating a museum as a tourist product using the SERVQUAL method. *Museum Management and Curatorship*, 20(3), 235-250.
- Nunnally, J. C. (1978). *Psychometric theory* (2e éd.). New York: McGraw-Hill.

- O'Muircheartaigh, C. A., Gaskell, G. D., & Wright, D. B. (1995). Weighing anchors: verbal and numeric labels for response scales. *Journal of Official Statistics*, 11(3), 295-307.
- O'Loughlin, C., & Coenders, G. (2004). Estimation of the European Customer Satisfaction Index: maximum likelihood versus partial least squares. Application to postal services. *Total Quality Management*, 15(9-10), 1231-1255.
- Oliver, R. L. (1977). Effect of expectations and disconfirmation on post-exposure product evaluations: an alternative explanation. *Journal of Applied Psychology*, 62(4), 480-486.
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17(November), 460-469.
- Oliver, R. L. (1981). Measurement and evaluation of satisfaction process in retail settings. *Journal of Retailing*, 57, 25-48.
- Oliver, R. L. (1997). *Satisfaction: a behavioral perspective on the consumer*. New York, NY: McGraw Hill.
- Oliver, R. L., & Bearden, W. O. (1985). Crossover effects in the theory of reasoned action: a moderating influence attempt. *Journal of Consumer Research*, 12(3), 324-340.
- Osborne, J. W. (2008). *Best practices in quantitative methods*: Sage Publications.
- Ostrom, O., & Iacobucci, D. (1995). Consumer trade-offs and the evaluation of services. *Journal of Marketing*, 59, 17-28.
- Ostrom, T. M. (1987). Bipolar survey items: an information processing perspective. Dans H.-J. Hippler, N. Schwarz & S. Sudman (Éds.), *Social information processing and survey methodology* (pp. 71-85). New York, NY: Springer-Verlag.
- Oterhals, K., Hanestad, B. R., Eide, G. E., & Hanssen, T. A. (2006). The relationship between in-hospital information and patient satisfaction after acute myocardial infarction. *European Journal of Cardiovascular Nursing*, 5(4), 303-310.
- Page, T. J. J., & Spreng, R. A. (2002). Difference scores versus direct effects in service quality measurement. *Journal of Service Research*, 4(3), 184-194.
- Parker, C., & Mathews, B. P. (2001). Customer satisfaction: contrasting academic and consumers' interpretations. *Marketing Intelligence & Planning*, 19(1), 38-44.
- Peterson, R. (1997). A quantitative analysis of rating-scale response variability. *Marketing Letters*, 8(1), 9-21.
- Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, 21(2), 381.

- Peterson, R. A., & Wilson, W. R. (1992). Measuring consumer satisfaction: fact and artifact. *Journal of the Academy of Marketing Science*, 20(1), 61-71.
- Phillips, W. (2001). What visitors value. Dans R. Adams (Éd.), *Museum visitor services manual* (pp. 14-15). Washington, DC: American Association of Museums.
- Phoenix Strategic Perspectives Inc. (2006). *Recherche sur la satisfaction de la clientèle de Service Canada : enquête de référence 2006*. Ottawa, Ontario: réalisée pour Service Canada.
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1-15.
- Rand, J. (2001). What visitors want and need. Dans R. Adams (Éd.), *Museum visitor services manual* (pp. 13-14). Washington, DC: American Association of Museums.
- Reinecke, J., Schmidt, P., & Ajzen, I. (1996). Application of the theory of planned behavior to adolescents condom use: a panel study. *Journal of Applied Social Psychology*, 26(9), 749-772.
- Renaud, J.-S. (2005). *Validation du modèle de fidélisation du client de Oliver*. Mémoire de maîtrise, Université Laval, Québec.
- Rogers, M. L. (1998). *An exploration of factors affecting customer satisfaction with selected history museum stores*. Texas Tech University.
- 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.
- Rust, R. T., Inman, J. J., Jia, J., & Zahorik, A. (1999). What you don't know about customer-perceived quality: the role of customer expectation distribution. *Marketing Science*, 18(1), 77-92.
- Scarpello, V., & Campbell, J. P. (1983). Job satisfaction: are all the parts there? *Personnel Psychology*, 36(3), 577-600.
- Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology*, 29, 65-88.
- Schmidt, F., & Strickland, T. (1998). *Sondages sur la satisfaction des clients : outil de mesures communes*. Ottawa, ON: Centre canadien de gestion.
- Schwarz, N. (1990). What respondents learn from scales: the informative functions of response alternatives. *International Journal of Public Opinion Research*, 2(3), 274-285.

- Schwarz, N. (1996). *Cognition and communication: judgmental biases, research methods, and the logic of conversation*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schwarz, N., Grayson, C. E., & Knauper, B. (1998). Formal features of rating scales and the interpretation of question meaning. *International Journal of Public Opinion Research, 10*(2), 177-183.
- Schwarz, N., & Hippler, H.-J. (1987). What response scales may tell your respondents: informative functions of response alternatives. Dans H.-J. Hippler, N. Schwarz & S. Sudman (Éds.), *Social information processing and survey methodology* (pp. 163-178). New York, NY: Springer-Verlag.
- Schwarz, N., Knauper, B., Hippler, H.-J., Noelle-Neumann, E., & Clark, L. (1991). Rating scales: numeric values may change the meaning of scale labels. *The Public Opinion Quarterly, 55*(4), 570-582.
- Shaw, D. G., Huffman, M. D., & Haviland, M. G. (1987). Grouping continuous data in discrete intervals: information loss and recovery. *Journal of Educational Measurement, 24*(2), 167-173.
- Si, S. X., & Cullen, J. B. (1998). Response categories and potential cultural bias: effects of an explicit middle point in cross-cultural surveys. *International Journal of Organizational Analysis, 6*(3), 218-230.
- Smith, T. W. (1995). *Little things matter: a sampler of how differences in questionnaire format can affect survey responses*. Document présenté à: Annual meeting of the American Association for Public Opinion Research.
- Söderlund, M. (1998). Customer satisfaction and its consequences on consumer behaviour revisited: the impact of different levels of satisfaction on word-of-mouth, feed-back to supplier and loyalty. *International Journal of Service Industry Management, 9*(2), 169-188.
- Söderlund, M., & Vilgon, M. (1999). Customer satisfaction and links to customer profitability: an empirical examination of the association between attitudes and behavior, *SSE/EFI Working Paper Series in Business Administration*: Stockholm School of Economics.
- Solomon, M. R., Polegato, R., & Zaichkowsky, J. L. (1999). *Consumer behaviour: buying, having, and being*. Toronto: Pearson Prentice Hall.
- Spreng, R. A., MacKeinzie, S. B., & Olshavsky, R. W. (1996). A reexamination of the determinants of consumer satisfaction. *Journal of Marketing, 60*(3), 15-32.
- Stauss, B., & Neuhaus, P. (1997). The qualitative satisfaction model. *International Journal of Service Industry Management, 8*(3), 236-249.

- Stephenson, N. L., & Herman, J. (2000). Pain measurement: a comparison using horizontal and vertical visual analogue scales. *Applied Nursing Research, 13*(3), 157-158.
- Strack, F., & Martin, L. L. (1987). Thinking, judging, and communicating: a process account of context effects in attitude surveys. Dans H.-J. Hippler, N. Schwarz & S. Sudman (Éds.), *Social information processing and survey methodology* (pp. 123-148). New York, NY: Springer-Verlag.
- Suen, H. K. (2009). Hakstian-Whalen test to compare among k Cronbach alpha values.
- Swan, J. E., Trawick, F. I., & Carrol, M. G. (Éds.). (1980). *Satisfaction related to predictive, desired expectations: a field study*. Bloomington, IN: Indiana University.
- Szymanski, D. M., & Henard, D. H. (2001). Customer satisfaction: a meta-analysis of the empirical evidence. *Journal of the Academy of Marketing Science, 29*(1), 20-35.
- Tang, R., Shaw, W. M. J., & Vevea, J. L. (1999). Towards the identification of the optimal number of relevance categories. *Journal of the American Society for Information Science, 50*(3), 254-264.
- Thurstone, L. L., & Chave, E. J. (1929). *The measurement of attitudes*. Chicago, IL: University of Chicago Press.
- Tourangeau, R. (2003). Cognitive aspects of survey measurement and mismeasurement. *International Journal of Public Opinion Research, 15*(1), 3-7.
- Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order. *Public Opinion Quarterly, 68*(3), 368-393.
- Tourangeau, R., Couper, M. P., & Conrad, F. (2007). Color, labels, and interpretative heuristics for response scales. *Public Opinion Quarterly, 71*(1), 91-112.
- Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The psychology of survey response*. Cambridge, U.K.: Cambridge University Press.
- van Herk, H., Poortinga, Y. H., & Verhallen, T. M. M. (2004). Response styles in rating scales: evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology, 35*(3), 346-360.
- Vanhamme, J. (2002). La satisfaction des consommateurs spécifique à une transaction : définition, antécédents, mesures et modes. *Recherche et Applications en Marketing, 17*(2), 55-85.
- Vavra, T. G. (1997). *Improving your measurement of customer satisfaction*. Milwaukee, Wisconsin: ASQ Quality Press.
- Vocational Psychology Research. (2007). Minnesota Satisfaction Questionnaire. Accédé le 11/02, 2007, de <http://www.psych.umn.edu/psylabs/vpr/msqinf.htm>

- Voss, K. E., Stem, D. E., & Fotopoulos, S. (2000). A comment on the relationship between coefficient alpha and scale characteristics. *Marketing Letters*, 11(2), 177-191.
- Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98(2), 219-235.
- Weisberg, H. F. (1992). Measures of spread. Dans *Central tendency and variability* (pp. 46-74). Newbury Park, Calif.: Sage Publications.
- Weiss, D. M., Dawis, R. V., England, G. W., & Lofquist, L. H. (1967). *Manual for the Minnesota Satisfaction Questionnaire*. Minneapolis: University of Minnesota, Industrial Relations Center.
- Weng, L.-J. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement*, 64(6), 956-972.
- Westbrook, R. A. (1980). A rating scale for measuring product/service satisfaction. *Journal of Marketing*, 44(4), 68-52.
- Westbrook, R. A. (1987). Product/consumption-based affective responses and postpurchase processes. *Journal of Marketing Research*, 24(August), 258-270.
- Westbrook, R. A., & Oliver, R. L. (1981). Developing better measures of consumer satisfaction: some preliminary results. *Advances in Consumer Research*, 8, 94-99.
- Westbrook, R. A., & Oliver, R. L. (1991). The dimensionality of consumption emotion patterns and consumer satisfaction. *Journal of Consumer Research*, 18(1), 84-91.
- Westbrook, R. A., & Reilly, M. D. (1983). Value-percept disparity: an alternative to the disconfirmation theory of consumer satisfaction. *Advances in Consumer Research*, 256-261.
- White, C., & Yu, Y.-T. (2005). Satisfaction emotions and consumer behavioral intentions. *The Journal of Services Marketing*, 19(6/7), 411-420.
- Wildt, A. R., & Mazis, M. B. (1978). Determinants of scale response: label versus position. *Journal of Marketing Research*, 15(2), 261-267.
- Willis, G. B. (2004). Cognitive interviewing revisited: a useful technique, in theory? Dans S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin & E. Singer (Éds.), *Methods for testing and evaluating survey questionnaires* (pp. 23-43). Hoboken, NJ: John Wiley & Sons Inc.
- Wirtz, J. (2001). Improving the measurement of customer satisfaction: a test of three methods to reduce halo. *Managing Service Quality*, 11(2), 99-112.

- Wirtz, J., & Lee, M. C. (2003). An examination of the quality and context-specific applicability of commonly used customer satisfaction measures. *Journal of Service Research*, 5(4), 345-355.
- Woodruff, R. B., Cadotte, E. R., & Jenkins, R. L. (1983). Modeling consumer satisfaction processes using experience-based norms. *Journal of Marketing Research*, 20(August), 296-304.
- Worthington, C. (2005). Patient satisfaction with health care: recent theoretical developments and implications for evaluation practice. *Canadian Journal of Program Evaluation*, 20(3), 41-63.
- Wyatt, R. C., & Meyers, L. S. (1987). Psychometric properties of four 5-point Likert type response scales. *Educational and Psychological Measurement*, 47(1), 27-35.
- Yi, Y. (1990). A critical review of consumer satisfaction. Dans V. A. Zeithaml (Éd.), *Review of Marketing*. Chicago, IL: American Marketing Association.
- Yu, C. H. (2003). *An introduction to computing and interpreting Cronbach Coefficient Alpha in SAS*. Document présenté à: Proceedings of the 26th SAS User Groups International Conference.
- Yucelt, U. (2000). Marketing museums: an ampirical investigation among museum visitors. *Journal of Nonprofit & Public Sector Marketing*, 8(3), 3-13.
- Yuksel, A., & Rimmington, M. (1998). Customer-satisfaction measurement. *Cornell Hotel and Restaurant Administration Quarterly*, 39(6), 60-70.
- Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing*, 60(2), 31-46.

Annexes

Annexe 1 : Définitions de la satisfaction des clients

Tableau A1.1 Définitions les plus courantes du concept de satisfaction des clients

Auteur(s)	Année	Définition
Howard et Sheth	1969	<p>« Chez l'acheteur, l'impression (état cognitif) d'avoir été adéquatement ou non récompensé pour les sacrifices qu'il a faits » (p. 145).</p> <p>Traduction libre de : « The buyer's cognitive state of being adequately or inadequately rewarded for the sacrifices he has undergone » (p. 145).</p>
Hunt ^a	1977	<p>« Une sorte de recul pris par rapport à l'expérience de consommation pour l'évaluer... l'évaluation que l'expérience était au moins aussi bonne qu'elle devait l'être ».</p> <p>Traduction libre de : « A kind of stepping away from an experience and evaluating it... the evaluation rendered that the experience was at least as good as it was supposed to be » (p. 459).</p>
Czepiel et Rosenberg	1977	<p>« Une attitude complexe de nature évaluative ».</p> <p>Traduction libre de : « A complex evaluative attitude » (p. 406).</p>
Swan, Trawick et Carroll ^a	1980	<p>« Une évaluation consciente ou un jugement cognitif sur la performance relative du produit ou sur le degré auquel le produit convenait pour son usage/utilisation. Une autre dimension de la satisfaction est l'affect envers le produit ».</p> <p>Traduction libre de : « A conscious evaluation or cognitive judgment that the product has performed relatively well or poorly or that the product was suitable or unsuitable for its use/purpose. Another dimension of satisfaction involves affect of feelings toward the product » (p. 17).</p>
Westbrook	1980	<p>« Réfère au degré de favorabilité de l'évaluation subjective de l'individu concernant les différentes conséquences et expériences associées à l'utilisation du produit ».</p> <p>Traduction libre de : « Refers to the favorability of the individual's subjective evaluation of the various outcomes and experiences associated with using or consuming it (product) » (référant à Hunt, 1977) (p. 49).</p>

Oliver	1981	<p>« Une évaluation de la surprise inhérente à l'acquisition d'un produit et/ou à une expérience de consommation. Dans son essence, l'état psychologique résultant de la combinaison des sentiments du consommateur avant l'achat et de l'émotion du consommateur entourant la disconfirmation des attentes ».</p> <p>Traduction libre de : « An evaluation of the surprise inherent in a product acquisition and/or consumption experience. In essence, the summary psychological state resulting when the emotion surrounding disconfirmed expectations is coupled with the consumer's prior feelings about the consumption experience » (p. 27).</p>
Churchill et Surprenant	1982	<p>« Conceptuellement, le résultat de la comparaison faite par l'acheteur des gains et des coûts engendrés par l'objet de consommation par rapport à ce qu'il anticipait avant l'achat. Opérationnellement, la satisfaction est similaire à l'attitude au sens où elle peut être mesurée par la sommation des scores de satisfaction sur les divers attributs de l'objet de consommation ».</p> <p>Traduction libre de : « Conceptually, an outcome of purchase and use resulting from the buyer's comparison of the rewards and costs of the purchase relative to anticipated consequences. Operationally, similar to attitude in that it can be assessed as a summation of satisfactions with various attributes » (p. 493).</p>
Engel et Blackwell ^c	1982	<p>« L'évaluation que l'alternative choisie est consistante par rapport aux croyances entretenues au départ concernant cette alternative ».</p> <p>Traduction libre de : « An evaluation that the chosen alternative is consistent with prior beliefs with respect to that alternative » (p. 501).</p>
Bearden et Teel	1983	<p>« Est une fonction des attentes du consommateur opérationnalisées en croyances par rapport aux attributs du produit ».</p> <p>Traduction libre de « A function of consumer expectations operationalized as product attribute beliefs (Olson & Dover, 1979) and disconfirmation » (p. 22). Pas de définition conceptuelle.</p>

LaBarbera et Mazursky	1983	« Évaluation postachat ».
Traduction libre de : « Postpurchase evaluation ». Cite également la définition de Oliver (1981) (p. 394).		
Westbrook et Reilly ^a	1983	« Une réponse émotionnelle associée aux objets de consommation (<i>i.e.</i> , produits, services), au magasin, aux comportements de consommation, de même qu'au marché en général ».
Traduction libre de : « An emotional response to the experiences provided by and associated with particular products or services purchased, retail outlets, or even molar patterns of behavior such as shopping and buyer behavior, as well as the overall marketplace » (p. 256).		
« Une réponse émotionnelle déclenchée par un processus d'évaluation cognitif à l'intérieur duquel les perceptions (ou croyances) par rapport à un objet, une action ou une condition sont comparées avec les valeurs (ou besoins, désirs) de l'individu ».		
Traduction libre de : « An emotional response triggered by a cognitive evaluative process in which the perceptions of (or beliefs about) an object, action, or condition are compared to one's values (or needs, wants, desires) » (p. 258).		
Woodruff, Cadotte et Jenkins	1983	« Une réponse émotionnelle ».
Traduction libre de : « An emotional response » (p. 302).		
Day ^a	1984	« L'évaluation de l'objet de consommation... la réponse du consommateur face l'évaluation de la différence perçue entre les attentes entretenues avant l'achat (ou un autre point de comparaison) et la performance actuelle du produit ».
Traduction libre de : « The evaluative response to the current consumption event...the consumer's response in a particular consumption experience to the evaluation of the perceived discrepancy between prior expectations (or some other norm of performance) and the actual performance of the product perceived after its acquisition » (p. 496).		

Cadotte, Woodruff et Jenkins	1987	« Conceptualisée comme un sentiment développé à partir d'une évaluation de l'utilisation de l'objet de consommation ». Traduction libre de : « Conceptualized as a feeling developed from an evaluation of the use experience » (p. 305).
Westbrook	1987	« Un jugement global concernant l'utilisation ou la consommation d'un produit ». Traduction libre de : « Global evaluative judgment about product usage/consumption » (p. 260)
Tse et Wilton	1988	« La réponse du consommateur face à la différence perçue entre les attentes de départ (ou un autre point de comparaison) et la performance actuelle du produit telle perçue après sa consommation ». Traduction libre de : « The consumer's response to the evaluation of the perceived discrepancy between prior expectations (or some norm of performance) and the actual performance of the product as perceived after its consumption » (p. 204).
Dufer et Moulins ^b	1989	La satisfaction est un état interne qui accompagne la confirmation des aspirations relatives au projet de consommation, celles-ci intégrant les attentes développées au sujet des produits et les normes de performance attachées à la classe dont il relève.
Oliver et Swan	1989	« (Par rapport au vendeur) Une fonction de l'équité, de la préférence et de la disconfirmation ». Traduction libre de : « (With the salesperson) A function of fairness, preference, and disconfirmation » (pp. 28-29). Pas de définition conceptuelle.
Westbrook et Oliver	1991	« Un jugement, après le choix d'un objet de consommation, concernant une décision d'achat spécifique ». Traduction libre de : « A postchoice evaluative judgment concerning a specific purchase selection » (Day, 1984) (p. 84).
Fornell	1992	« Une évaluation globale postachat ». Traduction libre de : « An overall postpurchase evaluation » (p.11).

Oliver ^a	1992	<p>« La satisfaction est un phénomène qui résume ce qui se passe au plan des attributs. Elle coexiste avec d'autres émotions liées à la consommation ».</p> <p>Traduction libre de : « Satisfaction is a summary attribute phenomenon coexisting with other consumption emotions » (p. 242).</p>
Mano et Oliver	1993	<p>« (La satisfaction envers un produit) est un jugement postconsommation similaire à une attitude variant en intensité sur un continuum d'hédonisme ».</p> <p>Traduction libre de : « (Product satisfaction) is an attitude-like postconsumption evaluative judgment (Hunt 1977) varying along the hedonic continuum (Oliver 1989; Westbrook and Oliver 1991) » (p. 454).</p>
Evrard	1993	<p>Un état psychologique résultant du processus d'achat et de consommation (p. 54).</p>
Halstead, Hartman et Schmidt	1994	<p>« Une réponse affective liée à une transaction spécifique. Cette réponse résulte de la comparaison que fait le consommateur entre la performance du produit et un point de comparaison antérieur à l'achat ».</p> <p>Traduction libre de : « A transaction-specific affective response resulting from the customer's comparison of product performance to some prepurchase standard (e.g., Hunt 1977; Oliver 1989) » (p. 122).</p>
Anderson, Fornell et Lehmann	1994	<p>« La satisfaction cumulative est une évaluation globale basée sur l'ensemble de l'achat et de l'expérience de consommation d'un produit ou d'un service ».</p> <p>Traduction libre de : « Cumulative satisfaction is an overall evaluation based on the total purchase and consumption experience with a good or a service over time » (p. 54).</p>

Ostrom et Iacobucci	1995	<p>« La satisfaction (et l'insatisfaction) du consommateur est un jugement relatif qui prend en considération l'ensemble de ce que le consommateur a investi pour obtenir l'objet de consommation ».</p> <p>Traduction libre de : « Customer satisfaction/dissatisfaction is thought to be a relative judgement that takes into consideration both the qualities and the efforts borne by a customer to obtain that purchase » (p. 17).</p>
Bloemer et Kasper	1995	<p>« La satisfaction peut être définie comme le résultat d'une évaluation subjective à savoir si l'alternative choisie rencontre ou excède les attentes ».</p> <p>Traduction libre de : « Satisfaction can be defined as the outcome of the subjective evaluation that the chosen alternative meets or exceeds the expectations » (p. 314).</p>
Oliver	1997	<p>« La satisfaction réfère à la réponse associée au degré de contentement du consommateur. Elle correspond au jugement qu'une caractéristique d'un produit ou d'un service, ou le produit ou le service lui-même, a apporté (ou apporte) un degré agréable de contentement, ce qui inclut que le degré de contentement soit trop peu élevé ou trop élevé ».</p> <p>Traduction libre de : « The consumer's fulfillment response. It is a judgment that a product or service feature, or the product or service itself, provided (or is providing) a pleasurable level of consumption-related fulfillment, including levels of under- or overfulfillment » (p. 13).</p>
Plichon ^b	1998	<p>La satisfaction est un état affectif provenant d'un processus d'évaluation affectif et cognitif qui survient lors d'une transaction spécifique.</p>

^aCité dans Giese et Cote (2000), pp. 5-8

^bCité dans Vanhamme (2002), p. 58

^cCité dans Vavra (2007), p. 4

**Annexe 2 : Échelle de satisfaction des clients de Oliver
(1980, 1997) traduite en français par Vanhamme (2002)**

Tableau A2.1 Échelle de satisfaction des clients de Oliver (1980, 1997) traduite en français par Vanhamme (2002)

Entourez le chiffre qui correspond le plus à votre opinion.	Pas du tout d'accord				Tout à fait d'accord
1. Je suis satisfait de la voiture que j'ai achetée	1	2	3	4	5
2. Choisir cette voiture fut un bon choix	1	2	3	4	5
3. Si je devais refaire mon choix, je choisirais un autre modèle de voiture	1	2	3	4	5
4. Je me sens coupable d'avoir acheté cette voiture	1	2	3	4	5
5. Je ne suis pas content d'avoir acheté cette voiture	1	2	3	4	5
6. Acheter cette voiture était ce qu'il fallait faire	1	2	3	4	5

**Annexe 3 : Directives pour réaliser la collecte de données
au Musée de la civilisation**

Le travail d'enquêteur

Le rôle de l'enquêteur est simplement de solliciter les visiteurs qui correspondent à la **population cible** afin de recruter des répondants pour l'enquête. **Les répondants doivent compléter le questionnaire par eux-mêmes.** Si ces derniers posent des questions à l'enquêteur, il ne doit en aucun cas suggérer des réponses ou émettre ses opinions personnelles : il doit rester neutre.

La **population cible** est la suivante :

- **18 ans et plus**
- **Résidents de la province de Québec**
- **Capables de lire le français**

Directives pour l'enquête

- ✓ Sélectionnez les **visiteurs qui ont terminé** leur visite du Musée. La visite doit avoir eu lieu la journée même, et au moins une exposition doit avoir été vue.
- ✓ Dans la mesure du possible, sollicitez tous ceux qui font partie de la population cible (les employés du Musée ne sont pas admissibles). En cas de doute, sollicitez-les et évaluez s'ils font partie de la population cible. S'ils n'en font pas partie, expliquez-leur qu'ils ne sont pas admissibles à la présente étude car celle-ci s'intéresse spécifiquement aux résidents de la province de Québec âgés de 18 ans et plus.
- ✓ Lorsque vous solliciter les visiteurs :
 1. Demandez-leur s'ils ont terminé leur visite.
 2. Dites-leur que *dans le cadre d'une thèse de doctorat à l'Université Laval, on distribue un questionnaire portant sur la satisfaction des visiteurs envers les expositions du Musée.*
 3. Ajouter qu'en complétant ce questionnaire, ils auront la chance de gagner l'un des quatre certificats-cadeaux de 50\$ chez Archambault.
 4. Terminez en leur demandant s'ils acceptent de répondre au questionnaire.
- ✓ Respectez l'ordre de distribution des six versions du questionnaire. L'ordre est le suivant : version 1 jusqu'à version 6, puis on recommence.
- ✓ Dans le cas d'un couple / groupe, offrez le questionnaire à tous.
- ✓ Vérifiez que le questionnaire est **complété** lorsque le visiteur le remet. Seuls les répondants qui remettent un questionnaire complété ont droit à un coupon de participation au tirage. Dans le cas où un répondant

remettrait un questionnaire non complété, offrez-lui de le compléter afin qu'il puisse obtenir son coupon de participation.

- ✓ Notez, sur le questionnaire, la date et l'heure à laquelle la personne a répondu au questionnaire.
- ✓ *En cas de refus* d'une personne faisant partie de la population cible, notez-le et indiquez la raison dans la grille servant à **mesurer le taux de réponse**.
- ✓ *Dans le cas des personnes ne faisant pas partie de la population cible*, comptabilisez-les dans la grille servant à **mesurer le taux de réponse**.
- ✓ L'objectif : 20-25 questionnaires complétés par jour (4 heures).
- ✓ À la fin de la journée, finissez de remplir la grille servant à mesurer le taux de réponse, déposez les questionnaires dans la boîte prévue à cette effet et mettez les coupons de participation dans l'enveloppe.

**Annexe 4 : Attributs d'une exposition muséale associés à
la satisfaction des visiteurs selon la littérature récente**

Tableau A4.1 Attributs d'une exposition muséale associés à la satisfaction des visiteurs

Auteur(s)	Attributs ou dimensions
Huo & Miller (2007)	Trois dimensions : <ul style="list-style-type: none"> • Venue attributes (Facilities) • Exhibition experience • Staff services
Harrison & Shaw (2004)	Trois dimensions : <ul style="list-style-type: none"> • Exhibition experience : <ul style="list-style-type: none"> ○ Educational ○ Stimulating ○ Relevant ○ Authentic ○ Contemplative ○ Emotive • Staff services : <ul style="list-style-type: none"> ○ Accessible ○ Informative ○ Friendly • Venue attributes : <ul style="list-style-type: none"> ○ Accessible ○ Ease of movement ○ Functional ○ Clean ○ Comfortable ○ Safe ○ Aesthetic
Musée du Louvre (2006)	Quatre attributs : <ul style="list-style-type: none"> • Qualité de l'accueil • Confort de la visite • Présentation et informations fournies sur les œuvres • Informations données pour s'orienter

Meehan (2002)

Dix attributs pour chacun des trois musées évalués :

Scienceworks museum :

- Having fun
- Things in working order
- Hands on/interactive
- Clear information
- Children learn by doing
- Cleanliness of building
- Curiosity stimulated
- Interesting information
- Activities for children
- Value for money

Immigration museum :

- Clear information
- Interesting information
- Things in working order
- Seeing real objects
- Staff friendly/helpful
- Staff knowledgeable
- Learning/new/different
- Balanced information
- Feel comfortable/relaxed
- Cleanliness of building

Melbourne museum :

- Things in working order
 - Cleanliness of building
 - Relevant information
 - Staff well informed
 - Interesting information
 - Staff friendly/helpful
 - Curiosity stimulated
 - Opportunity to learn
 - Seeing real objects
 - Value for money
-

England (2003)

Deux dimensions :

- Information :
 - Information about what to do in Greater Gainesville
 - Information in planning my trip to see “Sue”
- Museum :
 - Staff availability
 - Staff friendliness
 - Ability to see the exhibit up close
 - Cleanliness of the exhibit area
 - Exhibit itself
 - Physical layout of exhibit area
 - Variety of things to do in exhibit area
 - Ability to get tickets or book group package
 - Information about the exhibit

Nowacki (2005)

Sept dimensions :

- Surrounding of the attraction :
 - External appearance of the buildings
 - Suitable car park
 - Easy access for the elderly and disabled
 - Reception/ticket office :
 - Friendly and sensitive personnel
 - Ticket office personnel supplies detailed information
 - Clear posters with individual tourist attraction elements marked
 - Accessibility of specific information on the tourist attraction
 - Accessibility of clear maps of the attraction and vicinity
 - Computerized System of Tourist Information
 - Warnings of possible problems and potential danger
 - Accessibility of information about forbidden and limited behavior
 - Literature and souvenirs selling area :
 - Interesting and high quality choice: literature, souvenirs, etc.
 - Reasonable prices of these articles
 - Presentation of the products
-

- Exhibition :
 - Interestingly arranged and attractive exhibition
 - Exhibition presenting a given problem or a series of ideas in a logical order
 - Exhibition presenting natural environment issues
 - Exhibition presenting historical events or the history of the building/vicinity
 - Exhibition personnel making contact with the visitors
 - Exhibition personnel providing competent information
 - Exhibition focusing on the important and unique qualities of the building/vicinity
 - Exhibition providing interesting information for children
 - Exhibition stimulating discussion with family and friends
 - Exhibition aesthetics
 - Signs facilitating orientation and movement
 - Using appropriate means for information transfer
 - Clarity of exhibitions, clear and complementary descriptions
 - Personnel appearance

 - Catering :
 - Decor
 - Suitable menu
 - Aesthetics of food service
 - Cleanliness and freshness

 - Toilets :
 - Proper toilets
 - Cleanliness of toilets

 - General impression :
 - Attraction encouraging visits to the vicinity
 - Attraction arrangement which allows its appreciation
-

Yucelt (2000)

3 dimensions :

- Staff personnel and tour guides :
 - Staff courtesy
 - Staff friendly
 - Staff service
 - Staff knowledge
 - Staff availability

- Collections and offerings :
 - Educational value
 - Interpretative labels/signs
 - Collection/exhibits
 - Entertaining activities
 - Printed materials
 - Audio-visual programs

- Physical structure and related amenities :
 - Rest room facilities
 - Seating arrangements
 - Parking adequacy
 - Directional signs
 - Admission price

Rand (2001)

Onze besoins :

- Comfort
- Orientation
- Welcome
- Enjoyment
- Socializing
- Respect
- Communication
- Learning
- Choice and control
- Challenge and confidence
- Revitalization

Phillips (2001)

Onze attributs :

- Qualité des expositions
 - Facilité de se stationner/d'avoir accès à un stationnement
 - Indications pour s'orienter facilement dès l'entrée au musée
-

-
- Accueil chaleureux
 - Toilettes propres et qui ne manquent de rien
 - Possibilité de s'asseoir confortablement à l'intérieur des salles d'exposition
 - Heures d'ouverture qui conviennent à mon horaire
 - Qualité du service de cafétéria
 - Disponibilité d'espaces tranquilles
 - « Recovery process »
 - Personnalisation de l'expérience

Hooper-Greenhill (2001) Deux dimensions :

- Informations pour aider le visiteur à se diriger dans le musée
- Informations sur les expositions

Johns & Clark (2001) Sept attributs :

- Rapport qualité/prix
- Restaurant
- Hygiène des toilettes et de la salle d'eau
- Qualité de service de la boutique du musée
- Qualité des produits offerts dans la boutique du musée
- Degré auquel les attentes du visiteur ont été rencontrées
- Qualité de service du musée en général

Korn (2001) Quatre dimensions :

- Staff availability
- Staff courtesy
- Exhibit maintenance
- Exhibit availability (no broken exhibits, crowding)

Boone & Britt (2001) Trois dimensions :

- Exhibits :
 - Number of exhibits
 - Number of exhibits that can be touched
 - Exhibits change frequently
 - Number of exhibits for adults
 - All exhibits are working
 - Explaining how the exhibits relate to the area
 - Use of computers in exhibits
 - Amount of written information at each exhibit
-

-
- Amenities :
 - Cleanliness
 - Number of seats for resting in the museum
 - Amount of seating in the food area
 - Quality of food
 - Signs directing visitors to exhibits
 - Explanation about what is included in admission to the museum

 - Staffing :
 - Quantity of staff available to explain about exhibits
 - Individuals working in the admission ticket booth supply correct information
 - Number of « live » demonstrations run by the museum staff

Allaire (2000)

Treize attributs :

- Accueil
 - Signalisation à l'intérieur du musée
 - Propreté des lieux
 - Confort offert aux visiteurs
 - Intérêt à l'égard de l'exposition
 - Présentation des œuvres et des objets dans les salles
 - Éclairage dans les salles
 - Lisibilité des textes dans les salles
 - Variété des sujets d'exposition
 - Visites commentées
 - Boutique
 - Café
 - Visite (dans son ensemble)
-

Groupe SECOR (1996) 30 attributs :

- Facilité d'accès au Musée
 - À l'arrivée, facilité pour repérer la billetterie
 - Le temps d'attente pour passer à la billetterie
 - L'accueil, l'amabilité dont le personnel a fait preuve
 - L'information sur la programmation reçue avant d'entrer au Musée
 - Dès l'entrée au Musée, la facilité pour connaître les activités offertes aujourd'hui
 - L'atmosphère dégagée par l'architecture et l'aménagement
 - La qualité de la signalisation pour se diriger dans le Musée (salles d'exposition, services,...)
 - La variété d'expositions offertes aujourd'hui
 - L'aspect surprenant des expositions
 - Les expositions m'ont fait vivre des émotions, elles m'ont touché personnellement
 - Les expositions m'ont permis d'apprendre des choses
 - Le niveau d'approfondissement dans le traitement des sujets d'expositions
 - Le Musée permet la découverte de la civilisation québécoise
 - Le Musée permet la découverte de d'autres civilisations
 - Mes possibilités d'interagir avec les objets (jouer, toucher, ...)
 - La présence d'activités pour intéresser et occuper les enfants
 - La fréquence des visites guidées offertes au Musée
 - L'offre d'activités d'animation (concerts, spectacles, ...)
 - L'ensemble des moyens (équipements, documents, publications) que le Musée a mis à ma disposition pour bien comprendre le contenu des expositions
 - La disponibilité du personnel qualifié pour répondre à mes questions sur les expositions
 - La clarté des textes d'exposition dans les salles (panneaux lisibles)
 - La possibilité d'organiser ma visite comme je le voulais
 - La présence de bancs à différents endroits pour me reposer
 - Le nombre de visiteurs dans les salles
 - Le respect du calme par les autres visiteurs
-

-
- L'entretien des toilettes
 - La discrétion des gardiens de sécurité
 - La possibilité de rencontrer des gens
 - Globalement, diriez-vous que votre expérience d'aujourd'hui au Musée de la civilisation a été : pas du tout, peu, assez, très satisfaisante, exceptionnelle

Daignault (2008)

Neuf attentes des visiteurs :

- Être surpris(e), que ma curiosité soit attisée
 - Me détendre, me divertir, me changer les idées
 - Passer un bon moment, dans la convivialité
 - Me cultiver, découvrir de nouvelles choses
 - Partager, discuter, échanger des idées
 - Découvrir ou faire découvrir un lieu
 - Participer à l'éducation des enfants
 - Me plonger dans l'ambiance d'un pays, d'une époque
 - Vivre une expérience hors du commun, un moment magique
-

**Annexe 5 : Questionnaire de l'étude qualitative
exploratoire**

Titre de la recherche : La mesure de la satisfaction du client



Bonjour !

Ce questionnaire fait partie du projet de thèse de M. Jean-Sébastien Renaud, étudiant au doctorat à la Faculté des sciences de l'éducation de l'Université Laval. M. Renaud est supervisé par M. Pierre Valois, Ph.D., professeur au même département. De plus, ce questionnaire permettra au Musée de la civilisation de connaître davantage ses visiteurs afin de mieux les servir. Il s'adresse aux **résidents de la province de Québec** âgés de **18 ans et plus** et prend de 5 à 10 minutes à remplir.

**VOTRE COLLABORATION EST PRÉCIEUSE POUR NOUS PERMETTRE DE RÉALISER
CETTE ÉTUDE ET NOUS VOUS EN REMERCIONS D'Y PARTICIPER.**

Confidentialité

Les chercheurs et le Musée de la civilisation s'engagent envers vous à respecter l'anonymat de vos réponses. Vous pouvez donc répondre en toute confiance. Seuls les chercheurs et le Musée de la civilisation auront accès aux réponses, mais sans jamais connaître les noms des répondants. À cette fin, nous vous demandons de ne pas inscrire votre nom sur le questionnaire.

Participation et consentement

Vous êtes entièrement libre de répondre ou non à ce questionnaire. En retournant ce questionnaire, vous consentez à ce que les chercheurs associés au projet et le Musée de la civilisation de Québec utilisent vos réponses (de manière anonyme).

*SVP
RETOURNEZ CE QUESTIONNAIRE
À L'ENQUÊTEUR.*

Projet approuvé par le Comité d'éthique de la recherche avec des êtres humains de l'Université Laval (**no d'approbation 2008-207 A-1**), le 29 août 2008

1. Les gens qui vont voir une exposition dans un musée en ressortent avec une satisfaction ou une insatisfaction plus ou moins grande. Dans votre cas, qu'est-ce qui fait que vous ressortez satisfait ou insatisfait d'une exposition dans un musée?

a) Ce qui fait que vous ressortez SATISFAIT d'une exposition dans un musée?

b) Ce qui fait que vous ressortez INSATISFAIT d'une exposition dans un musée?

2. Êtes-vous déjà venu au *Musée de la civilisation* de Québec?

- 0. Non
- 1. Oui → En moyenne, combien de fois par année venez-vous au Musée de la civilisation?
 - a) Moins d'une fois par année
 - b) Environ une fois par année
 - c) Plus d'une fois par année

3. Avez-vous déjà visité d'autres musées que le *Musée de la civilisation* de Québec?

- 0. Non
- 1. Oui → En moyenne, combien de fois par année allez-vous visiter des musées (excluant le Musée de la civilisation de Québec)?
 - a) Moins d'une fois par année
 - b) Environ une fois par année
 - c) Plus d'une fois par année

4. Lors de votre visite d'aujourd'hui au *Musée de la civilisation* de Québec, combien d'expositions avez-vous visitées avant de remplir ce questionnaire?

5. De façon générale, votre visite d'aujourd'hui au *Musée de la civilisation* de Québec vous a :

- 1. Très satisfait(e)
- 2. Assez satisfait(e)
- 3. Peu satisfait(e)
- 4. Pas du tout satisfait(e)
- 99. Ne sais pas/Refus

6. Quel est votre sexe?

- 0. Féminin
- 1. Masculin

7. Auquel des groupes d'âge suivants appartenez-vous?

- 0. 18-24 ans
- 1. 25-34 ans
- 2. 35-44 ans
- 3. 45-54 ans
- 4. 55-64 ans
- 5. 65 ans et plus

8. Où est situé votre domicile principal? :

- 0. Région métropolitaine de Québec
- 1. Région métropolitaine de Montréal
- 2. Autre région du Québec
- 3. Autre province du Canada
- 4. Etats-Unis d'Amérique
- 5. France
- 6. Autre pays

9. Laquelle des occupations suivantes décrit le mieux votre situation actuelle?

- | | |
|--------------------------------------|----------------|
| 0. Travailleur(euse) à plein temps | 3. À la maison |
| 1. Travailleur(euse) à temps partiel | 4. Retraité(e) |
| 2. À la recherche d'un emploi | 5. Étudiant(e) |

10. Si vous êtes aux études, quel est votre niveau d'étude actuel?

- | | |
|------------------------------|-----------------------------------|
| 0. Primaire (7 ans et moins) | 2. Collégial (13 à 15 ans) |
| 1. Secondaire (8 à 12 ans) | 3. Universitaire (16 ans et plus) |

Si vous avez terminé vos études, quel est votre niveau d'étude complété?

- | | |
|------------------------------|-----------------------------------|
| 0. Primaire (7 ans et moins) | 2. Collégial (13 à 15 ans) |
| 1. Secondaire (8 à 12 ans) | 3. Universitaire (16 ans et plus) |

MERCI BEAUCOUP DE VOTRE PARTICIPATION !

**Annexe 6 : Réponses des participants au questionnaire de
l'étude qualitative exploratoire**

Ce qui fait que vous ressortez SATISFAIT d'une exposition dans un musée?

- apprendre quelque chose (connaissances nouvelles, découverte)
- interactif
- esthétisme de l'exposition
- exposition bien conçue (qualité des choses)
- présence d'activités spéciales
- bonne rotation des expositions
- si j'ai été ému, émerveillée (surprise), émotion
- objets beaux (beauté des choses présentées, qualité de l'art)
- objets intéressants (dépend des oeuvres présentés)
- panneaux instructifs
- vidéos instructifs (présence d'audiovisuels)
- descriptions des objets (textes explicatifs, niveau de langue, qualité des explications qui accompagnent les objets, clarté des explications)
- quantité d'objets (diversité des objets, beaucoup de choses à voir)
- origine des objets
- nouveauté
- organisation de l'information
- créativité dans l'aménagement de la salle d'exposition
- contenu de l'exposition
- salle d'exposition spatieuse
- information détaillée (beaucoup de choses à lire)
- apprendre quelque chose sur l'histoire (concerne l'histoire)
- apprendre quelque chose sur des façons de voir le passé, présent
- voir des choses qui nous rapprochent de notre pays
- voir de vieilles choses qui nous touchent
- lorsque j'ai vécu une "expérience en soi"
- l'exposition remplit ses promesses
- organisation spatiale bien faite (disposition des lieux, salle bien aménagée)
- ludique
- on peut toucher et voir
- la diversité
- l'exclusivité (l'originalité)
- thème qui m'intéresse (sujet traité et son niveau d'intérêt)
- apprendre sur les cultures (à caractère social)
- manière dont s'est présenté, l'aspect visuel
- objets plus vieux, antiques
- évasion dans l'histoire, dans le temps
- avoir le temps
- intéressant (passionnant, vs exposition ennuyante)
- possibilité de faire un visite explicative avec des écouteurs
- présentation des objets
- concerne notre propre histoire
- exposition propose une certaine réflexion (amène à aller plus loin dans ma réflexion, regard apporté sur le thème)
- pertinence de la thématique
- quantité d'informations (bonne information, information disponible, répond aux interrogations)
- pertinence des objets présentés dans le cadre de la thématique (cohérence de l'exposition)
- choses qui rappellent notre histoire, notre pays
- culturel (parfaire ma culture)
- bien présenté (présentation des objets)

- dépend des objets présentés (ce qu'on y montre)
- disposition des objets (présentation)
- convivialité de l'exposition
- ambiance (éclairage, ambiance sonore, musique)
- visuel avec texte à écouter pas trop long
- sujet relié à mes voyages
- possibilité de s'asseoir
- climatisation (température agréable)
- ça fait sortir
- ne sait pas
- plusieurs types de choses exposées: média, objets
- accessibilité de l'information
- guide dynamique (à l'écoute, passionné, empathique)
- lorsqu'il y a des trucs didactiques (didactique)
- Sujet d'actualité
- qualité de l'information (actuelle, qualité de la recherche)
- anecdotes
- paix, quiétude, pas pressé, silence, relaxant
- donne envie d'en savoir encore plus sur le sujet de l'exposition
- exposition plus moderne

Ce qui fait que vous ressortez INSATISFAIT d'une exposition dans un musée?

- pas d'ordre de circulation (ordre à suivre, chronologie, indications du circuit à suivre, pas de structure dans l'exposition)
- matériel exposé se retrouve au centre d'achat tout près (choses qu'on a pas l'habitude voir au quotidien, objets rares)
- si mes attentes sont déçues
- ne pas voir la pertinence ou l'intérêt de l'exposition
- salle d'exposition exigüe
- salle d'exposition qui donne le goût de sortir
- information compacte, serrée, tassée, écriture trop petite
- trop de matière à lire (trop d'écriture, quantité de lignes à lire, que du texte)
- achalandage trop élevé (affluence, pas trop de personnes dans le musée)
- aspect organisation
- lorsque rien n'attire mon attention (ne me sens pas interpellée par l'exposition)
- explications trop poussées
- manque d'atmosphère autour de l'information
- nombre de pièces trop grand = fatigue (exposition trop longue)
- objets trop moderne
- accueil froid
- bruits ambiants
- trop abstrait (vs voir des objets concrets)
- manque d'intérêt
- éclairage mauvais
- explications pas claires
- audiovisuels trop longs
- trop de vidéos (équilibre entre lectures, audio et vidéo; trop de films, de virtuel ou des représentations symboliques)
- superficiel
- mon incompréhension

**Annexe 7 : Questionnaire de l'étude qualitative
confirmatoire**

Titre de la recherche : La mesure de la satisfaction du client



Bonjour !

Ce questionnaire fait partie du projet de thèse de M. Jean-Sébastien Renaud, étudiant au doctorat à la Faculté des sciences de l'éducation de l'Université Laval. M. Renaud est supervisé par M. Pierre Valois, Ph.D., professeur au même département. Les résultats obtenus suite à la passation de ce questionnaire permettront au Musée de la civilisation de connaître davantage ses visiteurs afin de mieux les servir. Il s'adresse aux **résidents de la province de Québec âgés de 18 ans et plus** et prend 10 minutes à remplir.

**VOTRE COLLABORATION EST PRÉCIEUSE POUR NOUS PERMETTRE DE RÉALISER
CETTE ÉTUDE ET NOUS VOUS REMERCIONS D'Y PARTICIPER.**

Confidentialité

Les chercheurs et le Musée de la civilisation s'engagent envers vous à respecter l'anonymat de vos réponses. Vous pouvez donc répondre en toute confiance. Seuls les chercheurs et le Musée de la civilisation auront accès aux réponses, mais sans jamais connaître les noms des répondants. Aussi, nous vous demandons de ne pas inscrire votre nom sur le questionnaire.

Participation et consentement

Vous êtes entièrement libre de répondre ou non à ce questionnaire. En le retournant, nous considérerons que vous consentez à ce que les chercheurs associés au projet et le Musée de la civilisation utilisent vos réponses (de manière anonyme).

***SVP
RETOURNEZ CE QUESTIONNAIRE
À L'ENQUÊTEUR.***

Projet approuvé par le Comité d'éthique de la recherche avec des êtres humains de l'Université Laval (no d'approbation 2008-207 A-2), le 23 septembre 2008

4. Les gens qui vont voir une exposition dans un musée en ressortent avec une satisfaction ou une insatisfaction plus ou moins grande. Dans votre cas, habituellement qu'est-ce qui fait que vous ressortez satisfait ou insatisfait d'une exposition dans un musée?

a) Ce qui fait que vous ressortez SATISFAIT d'une exposition dans un musée ?

b) Ce qui fait que vous ressortez INSATISFAIT d'une exposition dans un musée ?

10. *...si vous êtes aux études*, quel est votre niveau d'étude actuel ?
- | | |
|------------------------------|-----------------------------------|
| 0. Primaire (7 ans et moins) | 2. Collégial (13 à 15 ans) |
| 1. Secondaire (8 à 12 ans) | 3. Universitaire (16 ans et plus) |
- ...si vous avez terminé vos études*, quel est votre niveau d'étude complété ?
- | | |
|------------------------------|-----------------------------------|
| 0. Primaire (7 ans et moins) | 2. Collégial (13 à 15 ans) |
| 1. Secondaire (8 à 12 ans) | 3. Universitaire (16 ans et plus) |

MERCI BEAUCOUP DE VOTRE PARTICIPATION !

**Annexe 8 : Première version de la mesure multi-items de
satisfaction (72 items)**

Tableau A8.1 Les 72 items de la première version de la mesure multi-items de satisfaction des visiteurs envers les expositions

<p>Considérant l'ensemble des expositions que vous avez vues aujourd'hui au Musée de la civilisation, quel est votre degré de satisfaction envers chacun des aspects suivants?</p>
L'apprentissage et la réflexion
1. La possibilité d'apprendre des choses intéressantes
2. Le degré auquel votre goût d'apprendre sur les thèmes traités a été stimulé
3. La possibilité de parfaire votre culture générale
4. La possibilité de découvrir des choses nouvelles
5. La possibilité d'apprendre sur l'histoire
6. La possibilité d'apprendre sur les civilisations
7. La possibilité d'apprendre sur le Québec
8. La qualité des réflexions proposées par les expositions
9. Les réflexions que suscitent chez vous les expositions
Le divertissement
10. Le plaisir à visiter les expositions
11. La possibilité de passer un bon moment
12. L'aspect relaxant de la visite des expositions
13. La capacité des expositions à vous impressionner, vous émerveiller
14. Le côté divertissant de la visite des expositions
15. La capacité des expositions à vous faire vivre des émotions
16. La capacité des expositions à vous toucher personnellement
17. La capacité des expositions à vous captiver
18. La possibilité de vivre une expérience hors du commun
19. La capacité des expositions à vous plonger dans une ambiance particulière
Les thèmes des expositions
20. Les thèmes des expositions
21. La capacité des expositions à traiter les thèmes de manière intéressante
22. La profondeur avec laquelle les thèmes des expositions sont traités
23. La capacité des expositions à toucher à plusieurs facettes des thèmes traités
24. La cohérence entre le thème et le contenu des expositions

Le contenu des expositions
25. La cohérence dans le contenu des expositions
26. La facilité à comprendre les expositions
27. L'originalité du contenu des expositions
28. La quantité d'objets présentés dans les expositions
29. La diversité du contenu dans les expositions
30. La quantité de contenu multimédia dans les expositions
31. L'équilibre entre le contenu multimédia et la quantité d'objets présentés dans les expositions
32. L'intérêt du contenu des expositions
33. La possibilité de voir des choses que j'ai rarement ou jamais l'occasion de voir
34. La possibilité de voir des objets historiques
35. La possibilité de voir de belles choses
36. La quantité d'explications qui accompagnent les objets exposés
37. La pertinence des explications qui accompagnent les objets exposés
38. La clarté des explications qui accompagne les objets exposés
39. La facilité à trouver rapidement l'information que vous cherchez sur les fiches explicatives
40. Les possibilités d'interagir dans les expositions (jouer, toucher, voir, entendre, sentir, etc.)
41. Le matériel interactif
42. Les présentations audiovisuelles
43. La fréquence des visites guidées
44. Les possibilités d'utiliser un audio-guide (visite avec des écouteurs)
Les salles d'exposition
45. La distance qui sépare les fiches explicatives des objets exposés
46. La lisibilité des fiches explicatives
47. L'éclairage des fiches explicatives
48. L'éclairage dans la salle d'exposition
49. L'éclairage des objets exposés
50. La grandeur des salles d'exposition
51. L'espace disponible pour contempler les objets exposés
52. L'espace disponible pour circuler dans les salles d'exposition
53. L'achalandage dans les salles d'exposition
54. Le respect du calme par les autres visiteurs
55. La possibilité de faire une visite agréable malgré l'achalandage
56. La facilité à repérer le parcours de visite dans les expositions
57. La quantité d'indications pour vous aider à vous diriger dans les expositions

58. La manière de mettre en valeur les objets exposés
59. L'esthétisme des salles d'exposition
60. L'aménagement des salles d'exposition
61. La créativité dans l'aménagement des salles d'exposition
62. L'ambiance dégagée par l'architecture et l'aménagement des salles d'exposition
63. L'ambiance sonore créée par la musique dans les salles d'exposition
64. La température ambiante à l'intérieur des salles d'exposition
65. La propreté des salles d'exposition
66. La quantité de sièges mis à votre disposition pour vous reposer dans les salles d'exposition
Le personnel assigné aux salles d'expositions
67. La quantité de personnel disponible pour répondre à vos questions à l'intérieur des salles d'exposition
68. La capacité du personnel à l'intérieur des salles d'exposition à répondre à vos questions sur les expositions
69. L'apparence soignée du personnel dans les salles d'exposition
70. La courtoisie du personnel dans les salles d'exposition
71. Le professionnalisme du personnel dans les salles d'exposition
72. La discrétion des agents de sécurité présents dans les salles d'exposition

Échelle de réponse :

Très Insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très Satisfait
1	2	3	4	5

**Annexe 9 : Première version de la mesure de l'attitude à
l'égard de la fréquentation du Musée de la civilisation
(26 items)**

Pour chacun des énoncés suivants, veuillez SVP indiquer votre degré d'accord ou de désaccord à l'aide de l'échelle de réponse ci-dessous :

Complètement en désaccord	Assez en désaccord	Légèrement en désaccord	Légèrement d'accord	Assez d'accord	Complètement d'accord
1	2	3	4	5	6

- 1) Aller au MCQ est une activité que je trouve très plaisante
- 2) Aller au MCQ est une activité ennuyante*
- 3) Aller au MCQ est une activité que je trouve trop sérieuse*
- 4) Aller au MCQ est une activité que je trouve trop intellectuelle*
- 5) Une visite au MCQ requiert trop de temps*
- 6) Aller au MCQ est une bonne idée de sortie
- 7) Aller au MCQ est du temps de loisir bien investi.
- 8) Aller au MCQ est dispendieux*
- 9) Aller au MCQ est compliqué*
- 10) Aller au MCQ est relaxant
- 11) Aller au MCQ est pour moi un divertissement de qualité
- 12) Aller au MCQ me permet d'apprendre des choses intéressantes
- 13) Aller au MCQ me permet de voir des choses intéressantes
- 14) Aller au MCQ me permet de découvrir des choses nouvelles pour moi
- 15) Aller au MCQ est captivant
- 16) Aller au MCQ est stimulant
- 17) Aller au MCQ me permet de passer un bon moment
- 18) Aller au MCQ me permet de voir des expositions intéressantes
- 19) Aller au MCQ me permet de voir des expositions qui m'émerveillent
- 20) Aller au MCQ me permet de voir des expositions d'une grande qualité
- 21) Aller au MCQ me permet de voir des expositions bien conçues
- 22) Aller au MCQ me permet de voir des expositions sur des thèmes qui m'intéressent
- 23) Aller au MCQ me permet de voir des expositions de qualité
- 24) Aller au MCQ me permet de voir des expositions qui répondent mes attentes
- 25) Aller au MCQ me permet de voir des expositions dans un environnement convivial
- 26) Aller au MCQ est une perte de temps*

*Item reflétant une attitude défavorable eu égard à la fréquentation du Musée de la civilisation.

Annexe 10 : Questionnaire des entrevues cognitives

Titre de la recherche : La mesure de la satisfaction du client



Bonjour !

Ce questionnaire fait partie du projet de thèse de M. Jean-Sébastien Renaud, étudiant au doctorat à la Faculté des sciences de l'éducation de l'Université Laval. M. Renaud est supervisé par M. Pierre Valois, Ph.D., professeur au même département. Les résultats obtenus suite à la passation de ce questionnaire permettront au Musée de la civilisation de connaître davantage ses visiteurs afin de mieux les servir. Il s'adresse aux **résidents de la province de Québec âgés de 18 ans et plus** et prend environ 15 minutes à remplir.

VOTRE COLLABORATION EST PRÉCIEUSE POUR NOUS PERMETTRE DE RÉALISER CETTE ÉTUDE ET NOUS VOUS REMERCIONS D'Y PARTICIPER

Confidentialité

Les chercheurs et le Musée de la civilisation s'engagent envers vous à respecter l'anonymat de vos réponses. Vous pouvez donc répondre en toute confiance. Seuls les chercheurs et les responsables du Musée de la civilisation auront accès aux réponses, mais sans jamais connaître les noms des répondants. Aussi, nous vous demandons de ne pas inscrire votre nom sur le questionnaire.

Participation et consentement

Vous êtes entièrement libre de répondre ou non à ce questionnaire. En le retournant, nous considérerons que vous consentez à ce que les chercheurs associés au projet et le Musée de la civilisation utilisent vos réponses (de manière anonyme).

***SVP
RETOURNEZ CE QUESTIONNAIRE
À L'ENQUÊTEUR***

Projet approuvé par le Comité d'éthique de la recherche avec des êtres humains de l'Université Laval (no d'approbation 2008-207 A-2), le 23 septembre 2008

Votre satisfaction envers les expositions du Musée de la civilisation

Considérant l'ensemble des expositions que vous avez vues aujourd'hui au Musée de la civilisation, nous vous demandons d'indiquer ci-dessous votre degré de satisfaction envers un ensemble de facteurs.

FACTEUR 1 - L'apprentissage et la réflexion		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
Êtes-vous satisfait...						
1.	...de ce que vous avez appris dans les expositions	1	2	3	4	5
2.	...du degré auquel les expositions ont stimulé votre intérêt envers les thèmes traités	1	2	3	4	5
3.	...de l'apport des expositions à votre culture générale	1	2	3	4	5
4.	...de la quantité de choses nouvelles que vous avez découvertes	1	2	3	4	5
5.	...de ce que les expositions vous ont appris sur l'histoire	1	2	3	4	5
6.	...de ce que les expositions vous ont appris sur les civilisations	1	2	3	4	5
7.	...de ce que les expositions vous ont appris sur le Québec	1	2	3	4	5
8.	...des réflexions apportées par les expositions	1	2	3	4	5
FACTEUR 2 – Le divertissement		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
Êtes-vous satisfait...						
9.	...du plaisir que vous avez eu à visiter les expositions	1	2	3	4	5
10.	...de l'aspect relaxant de la visite des expositions	1	2	3	4	5
11.	...de l'aspect divertissant de la visite des expositions	1	2	3	4	5
12.	...du degré auquel les expositions vous ont donné des émotions	1	2	3	4	5
13.	...de l'aspect captivant des expositions	1	2	3	4	5
FACTEUR 3 - Les thèmes des expositions		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
Êtes-vous satisfait...						
14.	...des thèmes des expositions	1	2	3	4	5
15.	...de la manière dont les thèmes sont traités	1	2	3	4	5
16.	...du degré d'approfondissement des thèmes	1	2	3	4	5

FACTEUR 4 - Le contenu des expositions		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
Êtes-vous satisfait...						
17.	...de la quantité d'objets présentés dans les expositions	1	2	3	4	5
18.	...de la diversité des objets présentés dans les expositions	1	2	3	4	5
19.	...de la présence d'objets qui vous ont émerveillé	1	2	3	4	5
20.	...de la présence de belles choses à voir	1	2	3	4	5
21.	...de la présence d'objets que vous avez rarement ou jamais l'occasion de voir	1	2	3	4	5
		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
22.	...de la présence d'objets historiques	1	2	3	4	5
23.	...de l'équilibre entre les différents types de contenus (ex. : textes, présentations audiovisuelles, objets, interactifs, etc.)	1	2	3	4	5
24.	...de l'originalité du contenu des expositions	1	2	3	4	5
25.	...de la pertinence du contenu des expositions	1	2	3	4	5
26.	...des explications fournies dans les vignettes d'objet (i.e., les textes qui décrivent les objets)	1	2	3	4	5
27.	...de la facilité à trouver rapidement l'information que vous cherchez dans les vignettes d'objet	1	2	3	4	5
28.	...de la qualité des textes dans les expositions	1	2	3	4	5
29.	...du nombre de textes dans les expositions	1	2	3	4	5
30.	...de la qualité des interactifs (ex. : les jeux, le matériel informatique, etc.)	1	2	3	4	5
31.	...du nombre d'interactifs	1	2	3	4	5
32.	...de la qualité des présentations audiovisuelles					
33.	...du nombre de présentations audiovisuelles	1	2	3	4	5
34.	...de la fréquence des visites commentées	1	2	3	4	5
35.	...du nombre de fois où vous avez eu l'occasion d'utiliser un audio-guide (visite avec des écouteurs)	1	2	3	4	5
FACTEUR 5 - Les salles d'exposition		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
Êtes-vous satisfait...						
36.	...par rapport à la distance qui sépare les objets exposés de leur vignette d'objet (i.e., les textes qui décrivent les objets)	1	2	3	4	5
37.	...de l'éclairage des objets exposés	1	2	3	4	5

38.	...de la lisibilité des vignettes d'objet	1	2	3	4	5
39.	...de l'éclairage des vignettes d'objet	1	2	3	4	5
40.	...de la lisibilité des textes dans les expositions	1	2	3	4	5
41.	...de l'éclairage des textes dans les expositions	1	2	3	4	5
42.	...de l'éclairage en général dans les salles d'exposition	1	2	3	4	5
43.	...de l'espace disponible pour circuler dans les salles d'exposition	1	2	3	4	5
44.	...par rapport à l'achalandage dans les salles d'exposition	1	2	3	4	5
45.	...du calme par les autres visiteurs dans les salles d'exposition	1	2	3	4	5
46.	...de la facilité à repérer le parcours de visite dans les expositions	1	2	3	4	5
47.	...des indications pour vous aider à vous diriger dans les expositions	1	2	3	4	5
48.	...de la manière de mettre en valeur les objets exposés	1	2	3	4	5
49.	...de l'esthétisme des salles d'exposition	1	2	3	4	5
50.	...de l'aménagement des salles d'exposition	1	2	3	4	5
51.	...de la créativité dans l'aménagement des salles d'exposition	1	2	3	4	5
		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
52.	...de l'ambiance dégagée par l'architecture et l'aménagement des salles d'exposition	1	2	3	4	5
53.	...de l'ambiance sonore créée par la musique dans les salles d'exposition	1	2	3	4	5
54.	...de la température ambiante à l'intérieur des salles d'exposition	1	2	3	4	5
55.	...de la propreté des salles d'exposition	1	2	3	4	5
56.	...de la quantité de sièges mis à votre disposition pour vous reposer dans les salles d'exposition	1	2	3	4	5
57.	...de la qualité des sièges mis à votre disposition pour vous reposer dans les salles d'exposition	1	2	3	4	5
FACTEUR 6 - Le personnel assigné aux salles d'expositions		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
Êtes-vous satisfait...						
58.	...du nombre de personnes disponibles pour répondre à vos questions sur les expositions	1	2	3	4	5
59.	...de la compétence du personnel pour répondre à	1	2	3	4	5

	vos questions sur les expositions					
60.	...de l'apparence soignée du personnel dans les salles d'exposition	1	2	3	4	5
61.	...de la courtoisie du personnel dans les salles d'exposition	1	2	3	4	5
62.	...du professionnalisme du personnel dans les salles d'exposition	1	2	3	4	5
63.	...de la discrétion des agents de sécurité présents dans les salles d'exposition	1	2	3	4	5
VOTRE ÉVALUATION GLOBALE		Très insatisfait	Insatisfait	Autant satisfait qu'insatisfait	Satisfait	Très satisfait
64.	<i>Compte tenu de vos attentes de départ</i> envers les expositions, quel est votre degré de satisfaction?	1	2	3	4	5
65.	Si vous considérez l'ensemble des coûts associés à une visite au Musée de la civilisation (ex. : déplacement, stationnement, billet d'entrée, etc.), quel est votre degré de satisfaction envers le <i>rapport qualité/prix des expositions</i> que vous avez visitées aujourd'hui?	1	2	3	4	5
66.	<i>Globalement</i> , quel est votre degré de satisfaction envers les expositions que vous avez visitées aujourd'hui ?	1	2	3	4	5

Votre attitude envers aller au Musée de la civilisation

Pour chacun des énoncés suivants, veuillez SVP indiquer votre degré d'accord ou de désaccord à l'aide de l'échelle de réponse

ci-dessous :

Complètement en désaccord	Assez en désaccord	Légèrement en désaccord	Légèrement d'accord	Assez d'accord	Complètement d'accord
1	2	3	4	5	6

Pour moi, aller au Musée de la civilisation...

67.	... est une activité ennuyante	1	2	3	4	5	6
68.	... est une activité trop sérieuse	1	2	3	4	5	6
69.	... est une activité culturelle plaisante	1	2	3	4	5	6
70.	... requiert trop de temps	1	2	3	4	5	6
71.	... est un bon choix de sortie	1	2	3	4	5	6
72.	... est du temps de loisir bien investit	1	2	3	4	5	6
73.	... c'est compliqué	1	2	3	4	5	6
74.	... c'est relaxant	1	2	3	4	5	6
75.	... est un très bon divertissement	1	2	3	4	5	6
76.	... me permet de voir des choses intéressantes	1	2	3	4	5	6
77.	... me permet d'apprendre des choses intéressantes	1	2	3	4	5	6
78.	... me permet de passer un bon moment	1	2	3	4	5	6
79.	... c'est dispendieux	1	2	3	4	5	6

Vos intentions par rapport au Musée de la civilisation

Pour chacun des énoncés suivants, veuillez SVP indiquer votre degré d'accord ou de désaccord à l'aide de l'échelle de réponse ci-dessous :

Complètement en désaccord	Assez en désaccord	Légèrement en désaccord	Légèrement d'accord	Assez d'accord	Complètement d'accord
1	2	3	4	5	6

80.	Au cours du prochain mois, j'ai l'intention <i>de recommander</i> aux gens de mon entourage d'aller au Musée de la civilisation.	1	2	3	4	5	6
81.	Au cours du prochain mois, j'ai l'intention <i>d'encourager fortement</i> les gens de mon entourage d'aller au Musée de la civilisation.	1	2	3	4	5	6
82.	Au cours du prochain mois, j'ai l'intention <i>de faire des commentaires positifs</i> aux gens de mon entourage concernant le Musée de la civilisation.	1	2	3	4	5	6
83.	<i>J'ai l'intention</i> de revenir au Musée de la civilisation dans les six prochains mois	1	2	3	4	5	6
84.	<i>Je tiens fortement</i> à revenir au Musée de la civilisation dans les six prochains mois	1	2	3	4	5	6

92.	<p>Où est situé votre domicile principal ?</p> <table border="0"> <tr> <td>0. Région de la Capitale-Nationale</td> <td>4. États-Unis d'Amérique</td> </tr> <tr> <td>1. Région de Montréal</td> <td>5. France</td> </tr> <tr> <td>2. Autre région du Québec</td> <td>6. Autre pays</td> </tr> <tr> <td>3. Autre province du Canada</td> <td></td> </tr> </table>	0. Région de la Capitale-Nationale	4. États-Unis d'Amérique	1. Région de Montréal	5. France	2. Autre région du Québec	6. Autre pays	3. Autre province du Canada	
0. Région de la Capitale-Nationale	4. États-Unis d'Amérique								
1. Région de Montréal	5. France								
2. Autre région du Québec	6. Autre pays								
3. Autre province du Canada									
93.	<p>Laquelle des occupations suivantes décrit le mieux votre situation actuelle ?</p> <table border="0"> <tr> <td>0. Travailleur(euse) à plein temps</td> <td>3. À la maison</td> </tr> <tr> <td>1. Travailleur(euse) à temps partiel</td> <td>4. Retraité(e)</td> </tr> <tr> <td>2. À la recherche d'un emploi</td> <td>5. Étudiant(e)</td> </tr> </table>	0. Travailleur(euse) à plein temps	3. À la maison	1. Travailleur(euse) à temps partiel	4. Retraité(e)	2. À la recherche d'un emploi	5. Étudiant(e)		
0. Travailleur(euse) à plein temps	3. À la maison								
1. Travailleur(euse) à temps partiel	4. Retraité(e)								
2. À la recherche d'un emploi	5. Étudiant(e)								
94.	<p><i>...si vous êtes aux études, quel est votre niveau d'étude actuel ?</i></p> <table border="0"> <tr> <td>0. Primaire (7 ans et moins)</td> <td>2. Collégial (13 à 15 ans)</td> </tr> <tr> <td>1. Secondaire (8 à 12 ans)</td> <td>3. Universitaire (16 ans et plus)</td> </tr> </table> <p><i>...si vous avez terminé vos études, quel est votre niveau d'étude complété ?</i></p> <table border="0"> <tr> <td>0. Primaire (7 ans et moins)</td> <td>2. Collégial (13 à 15 ans)</td> </tr> <tr> <td>1. Secondaire (8 à 12 ans)</td> <td>3. Universitaire (16 ans et plus)</td> </tr> </table>	0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)	1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)	0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)	1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)
0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)								
1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)								
0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)								
1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)								
MERCI BEAUCOUP DE VOTRE PARTICIPATION !									

**Annexe 11 : Questionnaire utilisé pour la mise à
l'épreuve des qualités psychométriques des cinq outils de
mesure développés**

Titre de la recherche : La mesure de la satisfaction du client



Bonjour !

Ce questionnaire fait partie du projet de thèse de M. Jean-Sébastien Renaud, étudiant au doctorat à la Faculté des sciences de l'éducation de l'Université Laval. M. Renaud est supervisé par M. Pierre Valois, Ph.D., professeur au même département. Les résultats obtenus suite à la passation de ce questionnaire permettront au Musée de la civilisation de connaître davantage ses visiteurs afin de mieux les servir. Il s'adresse aux **résidents de la province de Québec âgés de 18 ans et plus** et prend environ 10 minutes à remplir.

VOTRE COLLABORATION EST PRÉCIEUSE POUR NOUS PERMETTRE DE RÉALISER CETTE ÉTUDE ET NOUS VOUS REMERCIONS D'Y PARTICIPER

Confidentialité

Les chercheurs et le Musée de la civilisation s'engagent envers vous à respecter l'anonymat de vos réponses. Vous pouvez donc répondre en toute confiance. Seuls les chercheurs et les responsables du Musée de la civilisation auront accès aux réponses, mais sans jamais connaître les noms des répondants. Aussi, nous vous demandons de ne pas inscrire votre nom sur le questionnaire.

Participation et consentement

Vous êtes entièrement libre de répondre ou non à ce questionnaire. En le retournant, nous considérerons que vous consentez à ce que les chercheurs associés au projet et le Musée de la civilisation utilisent vos réponses (de manière anonyme).

***SVP
RETOURNEZ CE QUESTIONNAIRE
À L'ENQUÊTEUR***

Projet approuvé par le Comité d'éthique de la recherche avec des êtres humains de l'Université Laval (no d'approbation 2008-207 A-2), le 23 septembre 2008

Votre satisfaction envers les expositions du Musée de la civilisation

Considérant l'ensemble des expositions que vous avez vues aujourd'hui au Musée de la civilisation, nous vous demandons d'indiquer ci-dessous votre degré de satisfaction envers un ensemble de facteurs.

FACTEUR 1 - L'apprentissage et la réflexion		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
Êtes-vous satisfait...							
1.	...de ce que vous avez appris dans les expositions	1	2	3	4	5	6
2.	...du degré auquel les expositions ont stimulé votre intérêt envers les thèmes traités	1	2	3	4	5	6
3.	...de l'apport des expositions à votre culture générale	1	2	3	4	5	6
4.	...des réflexions apportées par les expositions	1	2	3	4	5	6
FACTEUR 2 – Le divertissement		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
Êtes-vous satisfait...							
5.	...du plaisir que vous avez eu à visiter les expositions	1	2	3	4	5	6
6.	...de l'aspect relaxant de la visite des expositions	1	2	3	4	5	6
7.	...de l'aspect divertissant de la visite des expositions	1	2	3	4	5	6
FACTEUR 3 - Les thèmes des expositions		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
Êtes-vous satisfait...							
8.	...des thèmes des expositions	1	2	3	4	5	6
9.	...de la manière dont les thèmes sont traités	1	2	3	4	5	6
10.	...du degré d'approfondissement des thèmes	1	2	3	4	5	6
FACTEUR 4 - Le contenu des expositions		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
Êtes-vous satisfait...							
11.	...de la quantité d'objets présentés dans les expositions	1	2	3	4	5	6
12.	...de la diversité des objets présentés dans les expositions	1	2	3	4	5	6
13.	...du nombre d'objets qui vous ont impressionné	1	2	3	4	5	6
14.	...du nombre d'objets que vous avez rarement ou jamais l'occasion de voir ailleurs	1	2	3	4	5	6
15.	...de l'équilibre entre les différents types de contenus (ex. : textes, films, objets, jeux interactifs, etc.)	1	2	3	4	5	6
16.	...de l'originalité du contenu des expositions	1	2	3	4	5	6
17.	...de la pertinence du contenu des expositions	1	2	3	4	5	6
18.	...de la quantité d'informations dans les vignettes d'objet (i.e., les textes qui décrivent les objets)	1	2	3	4	5	6
19.	...de la qualité des textes dans les expositions	1	2	3	4	5	6
20.	...du nombre de textes dans les expositions	1	2	3	4	5	6

21.	...de la qualité des interactifs (ex. : les jeux, le matériel informatique, etc.)	1	2	3	4	5	6
22.	...du nombre d'interactifs (ex. : les jeux, le matériel informatique, etc.)	1	2	3	4	5	6
		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
23.	...de la qualité des présentations audiovisuelles (i.e., les films)	1	2	3	4	5	6
24.	...du nombre de présentations audiovisuelles (i.e., les films)	1	2	3	4	5	6
FACTEUR 5 - Les salles d'exposition		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
Êtes-vous satisfait...							
25.	...de la facilité à trouver les textes descriptifs qui vont avec les objets	1	2	3	4	5	6
26.	...de la lisibilité des textes en général (ex. : textes sur les murs, textes qui décrivent les objets, etc.)	1	2	3	4	5	6
27.	...de l'éclairage des textes en général (ex. : textes sur les murs, textes qui décrivent les objets, etc.)	1	2	3	4	5	6
28.	...de l'éclairage en général dans les salles d'exposition	1	2	3	4	5	6
29.	...de l'espace disponible pour circuler dans les salles d'exposition	1	2	3	4	5	6
30.	...par rapport à l'achalandage dans les salles d'exposition	1	2	3	4	5	6
31.	...du calme dans les salles d'exposition	1	2	3	4	5	6
32.	...des indications pour vous aider à vous diriger dans les expositions	1	2	3	4	5	6
33.	...de la manière de mettre en valeur les objets exposés	1	2	3	4	5	6
34.	...de l'esthétisme des salles d'exposition	1	2	3	4	5	6
35.	...de l'aménagement des salles d'exposition	1	2	3	4	5	6
36.	...de la créativité dans l'aménagement des salles d'exposition	1	2	3	4	5	6
37.	...de l'ambiance dégagée par l'architecture et l'aménagement des salles d'exposition	1	2	3	4	5	6
38.	...de l'ambiance sonore créée par la musique dans les salles d'exposition	1	2	3	4	5	6
39.	...de la température ambiante à l'intérieur des salles d'exposition	1	2	3	4	5	6
40.	...de la propreté des salles d'exposition	1	2	3	4	5	6
41.	...de la quantité de sièges mis à votre disposition pour vous reposer dans les salles d'exposition	1	2	3	4	5	6
FACTEUR 6 - Personnel dans les salles d'expositions		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
Êtes-vous satisfait...							
42.	...du nombre de personnes disponibles pour répondre à vos questions sur les expositions	1	2	3	4	5	6

43.	...de l'apparence soignée du personnel dans les salles d'exposition	1	2	3	4	5	6
44.	...du professionnalisme du personnel dans les salles d'exposition	1	2	3	4	5	6
45.	...de la discrétion des agents de sécurité présents dans les salles d'exposition	1	2	3	4	5	6
VOTRE ÉVALUATION GLOBALE		Très Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Très Satisfait
46.	<i>Compte tenu de vos attentes de départ</i> envers les expositions, quel est votre degré de satisfaction?	1	2	3	4	5	6
47.	Si vous considérez l'ensemble des coûts associés à une visite au Musée de la civilisation (ex. : déplacement, stationnement, billet d'entrée, etc.), quel est votre degré de satisfaction envers le <i>rapport qualité/prix des expositions</i> que vous avez visitées aujourd'hui?	1	2	3	4	5	6
48.	<i>Globalement</i> , quel est votre degré de satisfaction envers les expositions que vous avez visitées aujourd'hui ?	1	2	3	4	5	6

Votre attitude envers aller au Musée de la civilisation

Pour chacun des énoncés suivants, veuillez SVP indiquer votre degré d'accord ou de désaccord à l'aide de l'échelle de réponse ci-dessous :

Complètement en désaccord	Assez en désaccord	Légèrement en désaccord	Légèrement d'accord	Assez d'accord	Complètement d'accord
1	2	3	4	5	6

Pour moi, aller au Musée de la civilisation...

49.	... est une activité ennuyante	1	2	3	4	5	6
50.	... est une activité trop sérieuse	1	2	3	4	5	6
51.	... est une activité culturelle plaisante	1	2	3	4	5	6
52.	... requiert trop de temps	1	2	3	4	5	6
53.	... est un bon choix de sortie	1	2	3	4	5	6
54.	... est du temps de loisir bien investi	1	2	3	4	5	6
55.	... c'est compliqué (i.e., accessibilité, heures d'ouverture, etc.)	1	2	3	4	5	6
56.	... c'est relaxant	1	2	3	4	5	6
57.	... est un très bon divertissement	1	2	3	4	5	6
58.	... me permet de voir des choses intéressantes	1	2	3	4	5	6
59.	... me permet d'apprendre des choses intéressantes	1	2	3	4	5	6
60.	... me permet de passer un bon moment	1	2	3	4	5	6
61.	... c'est dispendieux	1	2	3	4	5	6

Vos intentions par rapport au Musée de la civilisation

Pour chacun des énoncés suivants, veuillez SVP indiquer votre degré d'accord ou de désaccord à l'aide de l'échelle de réponse ci-dessous :

Complètement en désaccord	Assez en désaccord	Légèrement en désaccord	Légèrement d'accord	Assez d'accord	Complètement d'accord
1	2	3	4	5	6

62.	Au cours du prochain mois, j'ai l'intention <i>de recommander</i> aux gens de mon entourage d'aller au Musée de la civilisation.	1	2	3	4	5	6																				
63.	Au cours du prochain mois, j'ai l'intention <i>de faire des commentaires positifs</i> aux gens de mon entourage concernant le Musée de la civilisation.	1	2	3	4	5	6																				
64.	<i>J'ai l'intention</i> de revenir au Musée de la civilisation dans les six prochains mois.	1	2	3	4	5	6																				
65.	<i>Je tiens fortement</i> à revenir au Musée de la civilisation dans les six prochains mois.	1	2	3	4	5	6																				
66.	En pourcentage, quelle est la <i>probabilité</i> que vous reveniez au Musée de la civilisation dans les six prochains mois ?	<div style="display: flex; align-items: center;"> 0 % <table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> <td style="width: 10%; height: 20px;"></td> </tr> <tr> <td>0-10 %</td> <td>11-20 %</td> <td>21-30 %</td> <td>31-40 %</td> <td>41-50 %</td> <td>51-60 %</td> <td>61-70 %</td> <td>71-80 %</td> <td>81-90 %</td> <td>91-100 %</td> </tr> </table> 100 % </div>																0-10 %	11-20 %	21-30 %	31-40 %	41-50 %	51-60 %	61-70 %	71-80 %	81-90 %	91-100 %
0-10 %	11-20 %	21-30 %	31-40 %	41-50 %	51-60 %	61-70 %	71-80 %	81-90 %	91-100 %																		
67.	Environ <i>combien de fois</i> avez-vous l'intention de revenir au Musée de la civilisation dans les six prochains mois ? _____ (si aucune fois, inscrire « 0 »)																										

Votre profil

68.	<p>Lors de votre visite d'aujourd'hui au Musée de la civilisation, quelles expositions avez-vous visité ?</p> <p><input type="checkbox"/> Or des Amériques <input type="checkbox"/> Le musée du quai Branly - Regards sur la diversité culturelle</p> <p><input type="checkbox"/> Temps libre <input type="checkbox"/> Trésors du temps de la Nouvelle-France – La collection du Musée Stewart</p> <p><input type="checkbox"/> Nous, les premières nations <input type="checkbox"/> Neurones en action (<i>Espace découverte</i>)</p> <p><input type="checkbox"/> Le Temps des Québécois <input type="checkbox"/> Tour de terre (<i>Espace découverte</i>)</p> <p><input type="checkbox"/> Territoires <input type="checkbox"/> Place au Moyen Âge (<i>Espace découverte</i>)</p> <p><input type="checkbox"/> Urbanopolis</p>
69.	<p>Avant aujourd'hui, étiez-vous déjà venu au Musée de la civilisation de Québec ?</p> <p>0. Non</p> <p>1. Oui → En excluant votre visite d'aujourd'hui, combien de fois êtes-vous venu au Musée de la civilisation? _____</p> <p>→ En excluant votre visite d'aujourd'hui, combien de fois êtes-vous venu au Musée de la civilisation dans les six derniers mois? _____</p> <p>→ En moyenne, combien de fois par année venez-vous au Musée de la civilisation ?</p> <p>a) Moins d'une fois par année</p> <p>b) Environ une fois par année</p> <p>c) Plus d'une fois par année</p>
70.	<p>Avez-vous déjà visité d'autres musées que le Musée de la civilisation de Québec?</p> <p>0. Non</p> <p>1. Oui → En moyenne, combien de fois par année visitez-vous des musées (excluant le Musée de la civilisation de Québec) ?</p> <p>a) Moins d'une fois par année</p> <p>b) Environ une fois par année</p> <p>c) Plus d'une fois par année</p>
71.	<p>Quel est votre sexe ?</p> <p>0. Féminin 1. Masculin</p>
72.	<p>Auquel des groupes d'âge suivants appartenez-vous ?</p> <p>0. 18-24 ans 2. 35-44 ans 4. 55-64 ans</p> <p>1. 25-34 ans 3. 45-54 ans 5. 65 ans et plus</p>
73.	<p>Où est situé votre domicile principal ?</p> <p>0. Région de la Capitale-Nationale 4. États-Unis d'Amérique</p> <p>1. Région de Montréal 5. France</p> <p>2. Autre région du Québec 6. Autre pays</p> <p>3. Autre province du Canada</p>

74.	<p>Laquelle des occupations suivantes décrit le mieux votre situation actuelle ?</p> <table> <tr> <td>0. Travailleur(euse) à plein temps</td> <td>3. À la maison</td> </tr> <tr> <td>1. Travailleur(euse) à temps partiel</td> <td>4. Retraité(e)</td> </tr> <tr> <td>2. À la recherche d'un emploi</td> <td>5. Étudiant(e)</td> </tr> </table>	0. Travailleur(euse) à plein temps	3. À la maison	1. Travailleur(euse) à temps partiel	4. Retraité(e)	2. À la recherche d'un emploi	5. Étudiant(e)		
0. Travailleur(euse) à plein temps	3. À la maison								
1. Travailleur(euse) à temps partiel	4. Retraité(e)								
2. À la recherche d'un emploi	5. Étudiant(e)								
75.	<p><i>...si vous avez terminé vos études</i>, quel est votre plus haut niveau d'étude complété ?</p> <table> <tr> <td>0. Primaire (7 ans et moins)</td> <td>2. Collégial (13 à 15 ans)</td> </tr> <tr> <td>1. Secondaire (8 à 12 ans)</td> <td>3. Universitaire (16 ans et plus)</td> </tr> </table> <p><i>...si vous êtes aux études</i>, quel est votre niveau d'étude actuel ?</p> <table> <tr> <td>0. Primaire (7 ans et moins)</td> <td>2. Collégial (13 à 15 ans)</td> </tr> <tr> <td>1. Secondaire (8 à 12 ans)</td> <td>3. Universitaire (16 ans et plus)</td> </tr> </table>	0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)	1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)	0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)	1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)
0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)								
1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)								
0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)								
1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)								

MERCI BEAUCOUP DE VOTRE PARTICIPATION !

Annexe 12 : Questionnaire de l'étude

Titre de la recherche : La mesure de la satisfaction du client



Bonjour !

Ce questionnaire fait partie du projet de thèse de M. Jean-Sébastien Renaud, étudiant au doctorat à la Faculté des sciences de l'éducation de l'Université Laval. M. Renaud est supervisé par M. Pierre Valois, Ph.D., professeur au même département. Les résultats obtenus suite à la passation de ce questionnaire permettront au Musée de la civilisation de connaître davantage ses visiteurs afin de mieux les servir. Il s'adresse aux **résidents de la province de Québec âgés de 18 ans et plus** et prend environ 10 minutes à remplir.

VOTRE COLLABORATION EST PRÉCIEUSE POUR NOUS PERMETTRE DE RÉALISER CETTE ÉTUDE ET NOUS VOUS REMERCIONS D'Y PARTICIPER

Confidentialité

Les chercheurs et le Musée de la civilisation s'engagent envers vous à respecter l'anonymat de vos réponses. Vous pouvez donc répondre en toute confiance. Seuls les chercheurs et les responsables du Musée de la civilisation auront accès aux réponses, mais sans jamais connaître le nom des répondants. Aussi, nous vous demandons de ne pas inscrire votre nom sur le questionnaire.

Participation et consentement

Vous êtes entièrement libre de répondre ou non à ce questionnaire. En le retournant, nous considérerons que vous consentez à ce que les chercheurs associés au projet et le Musée de la civilisation utilisent vos réponses (de manière anonyme).

***SVP
RETOURNEZ CE QUESTIONNAIRE
À L'ENQUÊTEUR***

Votre satisfaction envers les expositions du Musée de la civilisation

Considérant l'ensemble des expositions que vous avez vues aujourd'hui au Musée de la civilisation, nous vous demandons d'indiquer ci-dessous votre degré de satisfaction envers un ensemble de facteurs.

FACTEUR 1 - L'apprentissage et la réflexion		Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
Êtes-vous satisfait...							
1.	...de ce que vous avez appris dans les expositions	1	2	3	4	5	6
2.	...du degré auquel les expositions ont stimulé votre intérêt envers les thèmes traités	1	2	3	4	5	6
3.	...de l'apport des expositions à votre culture générale	1	2	3	4	5	6
4.	...des réflexions apportées par les expositions	1	2	3	4	5	6
FACTEUR 2 – Le divertissement		Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
Êtes-vous satisfait...							
5.	...du plaisir que vous avez eu à visiter les expositions	1	2	3	4	5	6
6.	...de l'aspect relaxant de la visite des expositions	1	2	3	4	5	6
7.	...de l'aspect divertissant de la visite des expositions	1	2	3	4	5	6
FACTEUR 3 - Les thèmes des expositions		Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
Êtes-vous satisfait...							
8.	...des thèmes des expositions	1	2	3	4	5	6
9.	...de la manière dont les thèmes sont traités	1	2	3	4	5	6
10.	...du degré d'approfondissement des thèmes	1	2	3	4	5	6
FACTEUR 4 - Le contenu des expositions		Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
Êtes-vous satisfait...							
11.	...de la quantité d'objets présentés dans les expositions	1	2	3	4	5	6

12.	...de la diversité des objets présentés dans les expositions	1	2	3	4	5	6
13.	...du nombre d'objets qui vous ont impressionné	1	2	3	4	5	6
14.	...du nombre d'objets que vous avez rarement ou jamais l'occasion de voir ailleurs	1	2	3	4	5	6
15.	...de l'équilibre entre les différents types de contenus (ex. : textes, films, objets, jeux interactifs, etc.)	1	2	3	4	5	6
16.	...de l'originalité du contenu des expositions	1	2	3	4	5	6
17.	...de la pertinence du contenu des expositions	1	2	3	4	5	6
18.	...de la quantité d'informations dans les vignettes d'objet (i.e., les textes qui décrivent les objets)	1	2	3	4	5	6
19.	...de la qualité des textes dans les expositions	1	2	3	4	5	6
20.	...du nombre de textes dans les expositions	1	2	3	4	5	6
21.	...de la qualité des interactifs (ex. : les jeux, le matériel informatique, etc.)	1	2	3	4	5	6
22.	...du nombre d'interactifs (ex. : les jeux, le matériel informatique, etc.)	1	2	3	4	5	6
23.	...de la qualité des présentations audiovisuelles (i.e., les films)	1	2	3	4	5	6
24.	...du nombre de présentations audiovisuelles (i.e., les films)	1	2	3	4	5	6

FACTEUR 5 - Les salles d'exposition		Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
Êtes-vous satisfait...							
25.	...de la facilité à trouver les textes descriptifs qui vont avec les objets	1	2	3	4	5	6
26.	...de la lisibilité des textes en général (ex. : textes sur les murs, textes qui décrivent les objets, etc.)	1	2	3	4	5	6
27.	...de l'éclairage des textes en général (ex. : textes sur les murs, textes qui décrivent les objets, etc.)	1	2	3	4	5	6
28.	...de l'éclairage en général dans les salles d'exposition	1	2	3	4	5	6
29.	...de l'espace disponible pour circuler dans les salles d'exposition	1	2	3	4	5	6
30.	...par rapport à l'achalandage dans les salles d'exposition	1	2	3	4	5	6
31.	...du calme dans les salles d'exposition	1	2	3	4	5	6
32.	...des indications pour vous aider à vous diriger dans les expositions	1	2	3	4	5	6
33.	...de la manière de mettre en valeur les objets exposés	1	2	3	4	5	6
34.	...de l'esthétisme des salles d'exposition	1	2	3	4	5	6
35.	...de l'aménagement des salles d'exposition	1	2	3	4	5	6
36.	...de la créativité dans l'aménagement des salles d'exposition	1	2	3	4	5	6
37.	...de l'ambiance dégagée par l'architecture et l'aménagement des salles d'exposition	1	2	3	4	5	6
38.	...de l'ambiance sonore créée par la musique dans les salles d'exposition	1	2	3	4	5	6
39.	...de la température ambiante à l'intérieur des salles d'exposition	1	2	3	4	5	6

40.	...de la propreté des salles d'exposition	1	2	3	4	5	6
41.	...de la quantité de sièges mis à votre disposition pour vous reposer dans les salles d'exposition	1	2	3	4	5	6
FACTEUR 6 - Personnel dans les salles d'expositions		Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
Êtes-vous satisfait...							
42.	...du nombre de personnes disponibles pour répondre à vos questions sur les expositions	1	2	3	4	5	6
43.	...de l'apparence soignée du personnel dans les salles d'exposition	1	2	3	4	5	6
44.	...du professionnalisme du personnel dans les salles d'exposition	1	2	3	4	5	6
45.	...de la discrétion des agents de sécurité présents dans les salles d'exposition	1	2	3	4	5	6

VOTRE ÉVALUATION GLOBALE		Extrêmement Insatisfait	Assez Insatisfait	Légèrement Insatisfait	Légèrement Satisfait	Assez Satisfait	Extrêmement Satisfait
46.	<i>Compte tenu de vos attentes de départ</i> envers les expositions, quel est votre degré de satisfaction?	1	2	3	4	5	6
47.	Si vous considérez l'ensemble des coûts associés à une visite au Musée de la civilisation (ex. : déplacement, stationnement, billet d'entrée, etc.), quel est votre degré de satisfaction envers le <i>rapport qualité/prix des expositions</i> que vous avez visitées aujourd'hui?	1	2	3	4	5	6
48.	<i>Globalement</i> , quel est votre degré de satisfaction envers les expositions que vous avez visitées aujourd'hui ?	1	2	3	4	5	6

Votre attitude envers aller au Musée de la civilisation

Pour chacun des énoncés suivants, veuillez SVP indiquer votre degré d'accord ou de désaccord à l'aide de l'échelle de réponse ci-dessous :

Complètement en désaccord	Assez en désaccord	Légèrement en désaccord	Légèrement d'accord	Assez d'accord	Complètement d'accord
1	2	3	4	5	6

Pour moi, aller au Musée de la civilisation...

49.	... est une activité ennuyante	1	2	3	4	5	6
50.	... est une activité culturelle plaisante	1	2	3	4	5	6
51.	... requiert trop de temps	1	2	3	4	5	6
52.	... est du temps de loisir bien investi	1	2	3	4	5	6
53.	... c'est compliqué (i.e., accessibilité, heures d'ouverture, etc.)	1	2	3	4	5	6
54.	... c'est relaxant	1	2	3	4	5	6
55.	... est un très bon divertissement	1	2	3	4	5	6
56.	... me permet de voir des choses intéressantes	1	2	3	4	5	6
57.	... me permet de passer un bon moment	1	2	3	4	5	6
58.	... c'est dispendieux	1	2	3	4	5	6

Vos intentions par rapport au Musée de la civilisation

Pour les énoncés 59 à 61, veuillez SVP indiquer votre degré d'accord ou de désaccord à l'aide de l'échelle de réponse ci-dessous :

Complètement en désaccord	Assez en désaccord	Légèrement en désaccord	Légèrement d'accord	Assez d'accord	Complètement d'accord
1	2	3	4	5	6

59.	Au cours du prochain mois, j'ai l'intention <i>de recommander</i> aux gens de mon entourage d'aller au Musée de la civilisation.	1	2	3	4	5	6					
60.	Au cours du prochain mois, j'ai l'intention <i>de faire des commentaires positifs</i> aux gens de mon entourage concernant le Musée de la civilisation.	1	2	3	4	5	6					
61.	<i>J'ai l'intention</i> de revenir au Musée de la civilisation dans les six prochains mois.	1	2	3	4	5	6					
62.	Environ <i>combien de fois</i> avez-vous l'intention de revenir au Musée de la civilisation dans les six prochains mois ? _____ (si aucune fois, inscrire « 0 »)											
63.	En pourcentage, quelle est la <i>probabilité</i> que vous reveniez au Musée de la civilisation dans les six prochains mois ?											
	0 %	0-10 %	11-20 %	21-30 %	31-40 %	41-50 %	51-60 %	61-70 %	71-80 %	81-90 %	91-100 %	100 %

Votre profil

64.	<p>Premièrement, cochez [<input checked="" type="checkbox"/>] les expositions que vous avez visitées aujourd'hui.</p> <p>Deuxièmement, encerclez l'exposition que vous avez visitée en dernier (ex. : <input checked="" type="checkbox"/>).</p> <table border="0" style="width: 100%;"> <tr> <td><input type="checkbox"/> Extraterrestres... et si c'était vrai?</td> <td><input type="checkbox"/> Regards sur la diversité culturelle</td> </tr> <tr> <td><input type="checkbox"/> Temps libre</td> <td><input type="checkbox"/> Trésors du temps de la Nouvelle-France</td> </tr> <tr> <td><input type="checkbox"/> Nous, les premières nations</td> <td><input type="checkbox"/> Neurones en action (<i>Espace découverte</i>)</td> </tr> <tr> <td><input type="checkbox"/> Le Temps des Québécois</td> <td><input type="checkbox"/> Tour de terre (<i>Espace découverte</i>)</td> </tr> <tr> <td><input type="checkbox"/> Territoires</td> <td><input type="checkbox"/> Place au Moyen Âge (<i>Espace découverte</i>)</td> </tr> <tr> <td><input type="checkbox"/> Urbanopolis</td> <td></td> </tr> </table>	<input type="checkbox"/> Extraterrestres... et si c'était vrai?	<input type="checkbox"/> Regards sur la diversité culturelle	<input type="checkbox"/> Temps libre	<input type="checkbox"/> Trésors du temps de la Nouvelle-France	<input type="checkbox"/> Nous, les premières nations	<input type="checkbox"/> Neurones en action (<i>Espace découverte</i>)	<input type="checkbox"/> Le Temps des Québécois	<input type="checkbox"/> Tour de terre (<i>Espace découverte</i>)	<input type="checkbox"/> Territoires	<input type="checkbox"/> Place au Moyen Âge (<i>Espace découverte</i>)	<input type="checkbox"/> Urbanopolis	
<input type="checkbox"/> Extraterrestres... et si c'était vrai?	<input type="checkbox"/> Regards sur la diversité culturelle												
<input type="checkbox"/> Temps libre	<input type="checkbox"/> Trésors du temps de la Nouvelle-France												
<input type="checkbox"/> Nous, les premières nations	<input type="checkbox"/> Neurones en action (<i>Espace découverte</i>)												
<input type="checkbox"/> Le Temps des Québécois	<input type="checkbox"/> Tour de terre (<i>Espace découverte</i>)												
<input type="checkbox"/> Territoires	<input type="checkbox"/> Place au Moyen Âge (<i>Espace découverte</i>)												
<input type="checkbox"/> Urbanopolis													
65.	<p>Avant aujourd'hui, étiez-vous déjà venu au Musée de la civilisation de Québec ?</p> <p>0. Non</p> <p>1. Oui → En excluant votre visite d'aujourd'hui, combien de fois êtes-vous venu au Musée de la civilisation? _____</p> <p>→ En excluant votre visite d'aujourd'hui, combien de fois êtes-vous venu au Musée de la civilisation dans les six derniers mois? _____</p> <p>→ En moyenne, combien de fois par année venez-vous au Musée de la civilisation ?</p> <p>a) Moins d'une fois par année</p> <p>b) Environ une fois par année</p> <p>c) Plus d'une fois par année</p>												
66.	<p>Avez-vous déjà visité d'autres musées que le Musée de la civilisation de Québec?</p> <p>0. Non</p> <p>1. Oui → En moyenne, combien de fois par année visitez-vous des musées (excluant le Musée de la civilisation de Québec) ?</p> <p>a) Moins d'une fois par année</p> <p>b) Environ une fois par année</p> <p>c) Plus d'une fois par année</p>												
67.	<p>Quel est votre sexe ?</p> <p>0. Féminin 1. Masculin</p>												
68.	<p>Auquel des groupes d'âge suivants appartenez-vous ?</p> <table border="0" style="width: 100%;"> <tr> <td>0. 18-24 ans</td> <td>2. 35-44 ans</td> <td>4. 55-64 ans</td> </tr> <tr> <td>1. 25-34 ans</td> <td>3. 45-54 ans</td> <td>5. 65 ans et plus</td> </tr> </table>	0. 18-24 ans	2. 35-44 ans	4. 55-64 ans	1. 25-34 ans	3. 45-54 ans	5. 65 ans et plus						
0. 18-24 ans	2. 35-44 ans	4. 55-64 ans											
1. 25-34 ans	3. 45-54 ans	5. 65 ans et plus											

69.	<p>Où est situé votre domicile principal ?</p> <table border="0"> <tr> <td>0. Région de la Capitale-Nationale</td> <td>4. États-Unis d'Amérique</td> </tr> <tr> <td>1. Région de Montréal</td> <td>5. France</td> </tr> <tr> <td>2. Autre région du Québec</td> <td>6. Autre pays</td> </tr> <tr> <td>3. Autre province du Canada</td> <td></td> </tr> </table>	0. Région de la Capitale-Nationale	4. États-Unis d'Amérique	1. Région de Montréal	5. France	2. Autre région du Québec	6. Autre pays	3. Autre province du Canada	
0. Région de la Capitale-Nationale	4. États-Unis d'Amérique								
1. Région de Montréal	5. France								
2. Autre région du Québec	6. Autre pays								
3. Autre province du Canada									
70.	<p>Laquelle des occupations suivantes décrit le mieux votre situation actuelle ?</p> <table border="0"> <tr> <td>0. Travailleur(euse) à plein temps</td> <td>3. À la maison</td> </tr> <tr> <td>1. Travailleur(euse) à temps partiel</td> <td>4. Retraité(e)</td> </tr> <tr> <td>2. À la recherche d'un emploi</td> <td>5. Étudiant(e)</td> </tr> </table>	0. Travailleur(euse) à plein temps	3. À la maison	1. Travailleur(euse) à temps partiel	4. Retraité(e)	2. À la recherche d'un emploi	5. Étudiant(e)		
0. Travailleur(euse) à plein temps	3. À la maison								
1. Travailleur(euse) à temps partiel	4. Retraité(e)								
2. À la recherche d'un emploi	5. Étudiant(e)								
71.	<p>Si vous avez terminé vos études, quel est votre plus haut niveau d'étude complété ?</p> <table border="0"> <tr> <td>0. Primaire (7 ans et moins)</td> <td>2. Collégial (13 à 15 ans)</td> </tr> <tr> <td>1. Secondaire (8 à 12 ans)</td> <td>3. Universitaire (16 ans et plus)</td> </tr> </table> <p>Si vous êtes aux études, quel est votre niveau d'étude actuel ?</p> <table border="0"> <tr> <td>0. Primaire (7 ans et moins)</td> <td>2. Collégial (13 à 15 ans)</td> </tr> <tr> <td>1. Secondaire (8 à 12 ans)</td> <td>3. Universitaire (16 ans et plus)</td> </tr> </table>	0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)	1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)	0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)	1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)
0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)								
1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)								
0. Primaire (7 ans et moins)	2. Collégial (13 à 15 ans)								
1. Secondaire (8 à 12 ans)	3. Universitaire (16 ans et plus)								
MERCI BEAUCOUP DE VOTRE PARTICIPATION !									

**Annexe 13 : Évaluation de la validité nomologique de la
mesure de la satisfaction des clients : indices
d'ajustement et coefficient de détermination du modèle
structurel**

Tableau A13.1 Indices d'ajustement et coefficient de détermination selon la polarité et l'étiquetage (échelles de réponse en 6 points)

Étiquettes	Polarité	<i>n</i>	NFI	GFI	AGFI	χ^2	RSMEA	R^2
						(valeur <i>p</i>)		
Oui	Bipolaire	141	0,98	0,99	0,92	0,13	0,10	0,39
	Unipolaire	152	0,99	1,00	0,96	0,28	0,03	0,38
Non	Bipolaire	154	0,99	1,00	0,98	0,43	0,00	0,22
	Unipolaire	153	0,99	0,99	0,92	0,11	0,10	0,58

Tableau A13.2 Indices d'ajustement et coefficient de détermination selon la polarité (échelles de réponse en 10 points)

Polarité	<i>n</i>	NFI	GFI	AGFI	χ^2	RSMEA	R^2
					(valeur <i>p</i>)		
Bipolaire	157	1,00	1,00	1,00	0,76	0,00	0,48
Unipolaire	150	1,00	1,00	1,00	0,97	0,00	0,39

Tableau A13.3 Indices d'ajustement et coefficient de détermination selon le nombre de points de l'échelle de réponse

Polarité	Nombre de points	<i>n</i>	NFI	GFI	AGFI	χ^2	RSMEA	R^2
						(valeur <i>p</i>)		
Bipolaire	6	154	0,99	1,00	0,98	0,43	0,00	0,22
	10	157	1,00	1,00	1,00	0,76	0,00	0,48
Unipolaire	6	153	0,99	0,99	0,92	0,11	0,10	0,58
	10	150	1,00	1,00	1,00	0,97	0,00	0,39