Université Paris-Dauphine
Centre de Recherche DMS ERMES (UMR CNRS 7088)
École doctorale de Dauphine

L'ambivalence des consommateurs

Proposition d'un nouvel outil de mesure

Thèse pour l'obtention du titre de docteur en sciences de gestion
Présentée et soutenue publiquement le 22 mai 2014 par

Alice Audrezet

JURY

Directeur de thèse :	Monsieur Christian PINSON Professeur Emérite à l'INSEAD
Rapporteurs :	Madame Amina BÉJI-BÉCHEUR Professeure à l'Université Paris-Est
	Monsieur Alain JOLIBERT Professeur Emérite à l'Université Pierre Mendès France, Grenoble & Professeur à l'INSEEC
Suffragants :	Monsieur Denis DARPY Professeur à l'Université Paris-Dauphine
	Monsieur Svein Ottar OLSEN Professeur à l'Université Arctique de Norvège, Tromsø

L'université n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse : ces opinions doivent être considérées comme propres à leur auteur.

Sur proposition du Conseil de l'Edogest et sur celle du Conseil scientifique de l'Université Paris-Dauphine, le Conseil d'administration en date du 21 janvier 2008 a approuvé la suppression des mentions pour les thèses soutenues au sein d'Edogest, et ceci à compter du 1er janvier 2008.

L'ambivalence des consommateurs : proposition d'un nouvel outil de mesure

Résumé

La littérature en méthodologie pointe un important problème lié à l'utilisation des différentiels sémantiques pour mesurer les évaluations globales des individus. Au centre de ces continuums bipolaires opposant une paire d'adjectifs antagonistes, sont agrégées de manière artificielle et fâcheuse deux types d'évaluations : les réactions indifférentes et les réactions ambivalentes. Or, la distinction entre ces deux types d'évaluation est un enjeu dans la mesure où elles recouvrent des réalités différentes. En effet, l'indifférence correspond à l'absence de réaction positive ou négative, tandis que l'ambivalence décrit une évaluation composée à la fois de réactions positives et négatives. Cette recherche démontre l'intérêt d'utiliser un nouvel outil de mesure issu de recherches en psychologie, l'*Evaluative Space Grid* (Larsen & al., 2009), pour saisir la part d'ambivalence contenue dans les évaluations globales de consommateurs. A l'aide d'un design mixte comportant cinq études empiriques, nous montrons que l'*Evaluative Space Grid* présente des propriétés psychométriques aussi bonnes que le différentiel sémantique tout en permettant de différencier les réactions indifférentes des réactions ambivalentes, ce que ne peut pas faire un différentiel sémantique. Ce travail de recherche constitue donc une contribution méthodologique importante qui devrait intéresser chercheurs et praticiens.

Mots-clés : ambivalence des consommateurs, mesure des évaluations globales, différentiel sémantique, *Evaluative Space Grid*

Consumer Ambivalence: Proposition of a new Measurement Tool

Abstract

Literature on methodology reveals a serious problem related to the use of semantic differential scales to measure individuals' global evaluations. Two types of evaluations are incorporated—in an artificial and unsuitable way—into these bipolar continuums that contain two opposing adjectives: indifferent reactions and ambivalent reactions. The distinction between these two types of evaluation is significant because they reflect different realities. While indifference corresponds to an absence of positive or negative reactions, an ambivalent evaluation is composed of both positive and negative reactions. This research demonstrates the advantages of using a new measurement tool, the *Evaluative Space Grid* (Larsen & al., 2009), developed in psychology research; this tool is able to capture the ambivalent component in consumers' global evaluations. Using a mixed design method comprising five empirical studies, we demonstrate that the *Evaluative Space Grid* has psychometric properties that match those of semantic differential scales; however, in contrast to semantic differential scales, the *Evaluative Space Grid* is capable of differentiating between indifferent reactions and ambivalent reactions. This research therefore constitutes a major methodological contribution that will be of interest to both researchers and practitioners.

Key words: consumer ambivalence, global evaluation measurement, semantic differential scales, *Evaluative Space Grid*

Un instant après, Alice était à la poursuite du Lapin dans le terrier, sans songer comment elle en sortirait.

Pendant un bout de chemin le trou allait tout droit comme un tunnel, puis tout à coup il plongeait perpendiculairement d'une façon si brusque qu'Alice se sentit tomber comme dans un puits d'une grande profondeur, avant même d'avoir pensé à se retenir.

De deux choses l'une, ou le puits était vraiment bien profond, ou elle tombait bien doucement ; car elle eut tout le loisir, dans sa chute, de regarder autour d'elle et de se demander avec étonnement ce qu'elle allait devenir. D'abord elle regarda dans le fond du trou pour savoir où elle allait ; mais il y faisait bien trop sombre pour y rien voir. Ensuite elle porta les yeux sur les parois du puits, et s'aperçut qu'elles étaient garnies d'armoires et d'étagères ; çà et là, elle vit pendues à des clous des cartes géographiques et des images. […]

« Certes, » dit Alice, « après une chute pareille je me moquerai pas mal de dégringoler l'escalier ! Comme ils vont me trouver brave chez nous ! Je tomberais du haut des toits que je ne ferais pas entendre une plainte. »

Lewis Carroll, « Alice au pays des merveilles », Macmillan, 1869 (p.3-4)
Traduction par Henri Bué

– Remerciements –

« Réveillez-vous, chère Alice ! » lui dit sa sœur. « Quel long somme vous venez de faire ! »

« Oh ! J'ai fait un si drôle de rêve » dit Alice.

Et elle raconta […].

Lewis Carroll, « Alice au pays des merveilles », Macmillan, 1869 (p.192-193)
Traduction par Henri Bué

Au début de l'histoire, il y a un lapin blanc. Le mien n'était pas en retard, mais s'est avéré tout aussi passionnant.

Monsieur Pinson, vous avez changé ma vie. Parce que vous m'avez transmis la passion de la recherche et formé à la rigueur scientifique, parce que vous m'avez fait profiter, si généreusement de vos connaissances et de vos idées, je vous remercie très sincèrement. Merci de m'avoir ouvert les portes de ce métier que j'aime tant.

Je souhaite ensuite témoigner ma profonde reconnaissance aux Professeurs Amina Béji-Bécheur et Denis Darpy, dont les critiques constructives et bienveillantes formulées à l'occasion de la pré-soutenance m'ont aidée à positionner ma thèse et à prendre confiance en ma recherche. Je remercie également le Professeur Alain Jolibert de me faire l'honneur d'évaluer ce travail. Enfin, la participation de Svein Ottar Olsen au jury, revêt une dimension hautement symbolique à mes yeux. Svein Ottar, merci de m'avoir aidée, durant mon

septentrional séjour de recherche, à faire le choix de l'ambivalence. C'est une joie immense pour moi que tu aies accepté de venir du « Pôle Nord » pour en apprécier l'aboutissement.

Ma gratitude s'adresse ensuite à toute l'équipe du DMS. Je pense en particulier à Pierre Desmet, Pierre Volle et Guillaume Johnson qui lors du dernier séminaire doctoral m'ont assuré que, oui oui, c'était bien fini. Je remercie également Denis Guiot d'avoir patiemment écouté et finement commenté huit présentations successives sur le biais de négativité, alors que ma thèse traite finalement de l'ambivalence. Monsieur Guiot, j'espère que vous ne m'en tiendrez pas rigueur. Par ailleurs, et dans le but d'assurer mon éternelle participation au « wine & cheese », *social event* le plus envié de la recherche de gestion, je remercie très chaleureusement les exceptionnelles maîtres de conférences qui le coordonnent avec *maestria*. Il est finalement impossible d'évoquer le DMS sans remercier Astrid et Juliette pour leur diligence et leur extrême gentillesse.

Quelques « personnes morales » ont également compté. J'ai déjà évoqué Dauphine et le DMS, mais j'éprouve aussi une profonde gratitude envers l'ISG qui m'a permis de vivre une fin de thèse confortable, à l'abri de la précarité. Je tiens par ailleurs à exprimer ma reconnaissance envers la Chaire AFNOR Performance des Organisations, dirigée par Sylvie Rolland, pour son soutien financier déterminant dans la réalisation de ce travail doctoral. Rien n'aurait enfin été possible, sans les collectes de données (parfois loupées) au « labo d'expé ». Je remercie donc Denis Darpy d'en avoir porté le projet, ainsi que Maggie et Ziyed, pour avoir pris en charge sa délicate logistique.

J'en profite pour remercier les étudiants de Dauphine et de l'ISG, ainsi que tous les autres anonymes qui ont apporté leurs réponses à mes nombreux questionnaires. Merci aussi aux membres de ma famille et à mes amis qui ont inlassablement pré-testé ou fait suivre les liens de mes enquêtes.

Mes remerciements s'adressent ensuite aux Experts qui se sont mobilisés sur le dossier : tout d'abord à mes collègues chercheurs qui ont si joliment mis en mot (et parfois en image) la diagonale de la grille ; ensuite aux professionnels qui ont pris le temps de se pencher sur ce nouvel outil : Pierre Gomy (Millward Brown), Gilles Granger (Vinivi), Séverine Philardeau (TripAdvisor), Helen Zeitoun (GfK) et ma copine « Julie » qui se reconnaîtra. J'en profite pour souligner la discrète contribution de Dominique Mahut, ingénieur d'études statistiques au CNRS qui a su comprendre mon travail et y apporter d'intelligentes solutions.

Ces remerciements me donnent enfin l'occasion d'exprimer ma plus sincère reconnaissance envers l'inépuisable équipe de relecteurs critiques et géniaux qui ont permis d'enrichir, ou plus souvent d'alléger, ce document. Mille mercis à Caroline Ardelet, Celine Berrier-Lucas, Vivien Blanchet, Julien Bourjot, Romain Cadario, Audrey Davoust, Gwarlann de Kerviler, Ikram Nasr, Philippe Mérigot, Béatrice Parguel, Hélène Peton, Narjes Sassi, et bien sûr ma mère et ma tante Laurence… et de nombreuses fois Christian Pinson.
Si l'usage du « nous d'auteur » m'a parfois semblé artificiel, il prend pourtant tout son sens à l'évocation de cette interminable liste.

Du fond du cœur, je souhaite remercier la *team* BnF. Parce que pour finir une thèse, il ne faut pas juste de la motivation et du travail, mais aussi des tuiles aux amandes et des sms d'encouragement le matin, ma reconnaissance va tout spécialement à Claudine. Mention spéciale à Oussama qui a même accompagné ma fantaisie jusqu'à la file d'attente de la BPI. A partir d'aujourd'hui les amis, l'objet de nos mails ne sera plus seulement « BnF demain ? »

Mes remerciements vont également à tous les occupants de la A105 et de la A107 sans lesquels l'aventure aurait été tellement moins palpitante : Gwarlann et Caroline, les sœurs que je n'ai jamais eues ; Hélène (Daniel Morin et Nicolas Demorand) ; Romain, mon coach optimisateur – ce qui est chouette quand on fait sa thèse en deux ans, c'est que ça laisse pas mal de temps pour aider les copains – ; Vivien ma bonne étoile, Béa (mon modèle ?), Audrey, Julia et le body tonic ; et enfin, Philippe qui, avec la neutralité bienveillante d'un Jérôme Garcin, a regardé la chose se faire de loin.

Je tiens aussi à témoigner toute ma reconnaissance à mes incroyables collègues de l'ISG et en particulier à ceux de la très *select* « caserne » : Hicham-et-Ikram, l'autre Alice et Emmanuel, merci de m'avoir intégrée à votre « clan ». Par ailleurs, j'éprouve une profonde gratitude envers Narjes, qui en plus de m'avoir initiée à Lisrel avec une infinie patience, m'a conté la Tunisie : Sousse, Ksibet, Sfax et sa tante Catherine… Narjes, tu es la Shérazade des statistiques !

On insiste souvent sur l'importance du réseau dans le milieu de la recherche. C'est en voulant étendre le mien que j'ai rencontré Jérémy, l'antisystème. Et nous avons continué à « réseauter » en d'autres lieux. D'abord avec la geekissime Celine et quelques autres, puis avec une sacrée bande de copains. *#GroupeduVendredi*, je vous dédie cette thèse, en souvenir des typologies qui ne peuvent être débattues qu'au Supercoin.

Il y a aussi les amies de toujours, Jennifer, Iris et Emilie qui ont pardonné mon manque de temps et se sont pliées à mes plannings hallucinants. Je garderai un souvenir ému des pique-niques de l'été 2013 le long de la Seine ou au bord de la Fontaine Stravinsky.

J'en profite pour faire un clin d'œil à Régis qui a épaulé les débuts de mes errements.

Et j'en viens à Suzy, ma pauvre Suzy… Moi qui ai longtemps pensé occuper une place de choix dans tes souvenirs erasmussiens, aux côtés de Pieter-Jan ou à l'arrière de ton vélo. Me voilà terrifiée à l'idée que tout cela soit balayé par les écrasantes tours de la BnF et les salades à moins de 120 calories. Courage Suzy, toi aussi tu as presque fini !

Enfin, mes pensées se tournent vers Julien qui, comme Denver le dernier Dinosaure, est mon petit frère et bien plus encore. Merci de tout mon cœur Julien pour ton soutien psychologique, technique, humoristique, métaphysique, culturel et affectif… sans limite.

Et pour finir, mes remerciements les plus tendres vont à mes parents. Tout d'abord à ma mère, qui avec le soutien d'Albert, est indéniablement la relectrice la plus rapide de l'Ouest. Ensuite à mon père, le Edgard Hoover de la recherche en gestion. Qui d'autre que moi peut affirmer que son père connait les sujets de thèse de tous les doctorants du DMS ? Merci à vous deux pour votre inestimable implication.

– Sommaire –

INTRODUCTION **11**

CHAPITRE 1 - DEFINITION ET MESURE DE L'AMBIVALENCE **19**

SECTION 1. QU'EST-CE QUE L'AMBIVALENCE ? 19
SECTION 2. COMMENT MESURER L'AMBIVALENCE ? 26
SECTION 3. ETUDE PRELIMINAIRE SUR LA MESURE DES EVALUATIONS GLOBALES DE CONSOMMATEURS DANS LE E-COMMERCE 40

CHAPITRE 2 - L'*EVALUATIVE SPACE GRID*, UN OUTIL DE MESURE ALTERNATIF POUR SAISIR LA PART D'AMBIVALENCE DES CONSOMMATEURS **59**

SECTION 1. L'EVALUATIVE SPACE GRID 60
SECTION 2. APPLICATION DE L'*EVALUATIVE SPACE GRID* A LA MESURE DES EVALUATIONS DE CONSOMMATEURS 65

CHAPITRE 3 - VALIDITE CONVERGENTE DE L'*EVALUATIVE SPACE GRID* **79**

SECTION 1. ANALYSE DE LA VALIDITE CONVERGENTE DE LA GRILLE AVEC LES PAIRES DE CONTINUUMS UNIPOLAIRES (ETUDE 1) 79
SECTION 2. ANALYSE DE LA VALIDITE CONVERGENTE DE LA GRILLE AVEC LES DIFFERENTIELS SEMANTIQUES, LES DISCOURS DES CONSOMMATEURS ET UNE MESURE DE L'AMBIVALENCE SUBJECTIVE (ETUDE 2) 91

CHAPITRE 4 - ETUDE DE LA PERTINENCE DE L'*EVALUATIVE SPACE GRID* COMME OUTIL DE MESURE DE LA SATISFACTION CLIENT **103**

SECTION 1. PRESENTATION DE L'ETUDE ET FORMULATION DES HYPOTHESES 104
SECTION 2. METHODE 105
SECTION 3. ANALYSES ET RESULTATS 110
SECTION 4. CONCLUSION 127
SECTION 5. LIMITES 128

CHAPITRE 5 - PROCESSUS D'APPROPRIATION DE L'*EVALUATIVE SPACE GRID* PAR LES REPONDANTS : L'AVIS DES EXPERTS **133**

SECTION 1. QU'EN PENSENT LES CHERCHEURS ? 134
SECTION 2. QU'EN PENSENT LES PROFESSIONNELS ? 159
SECTION 3. CONCLUSION 167

CHAPITRE 6 - LA PLACE DE L'INCERTITUDE DANS LES EVALUATIONS EFFECTUEES AVEC L'*EVALUATIVE SPACE GRID* **171**

SECTION 1. CADRE THEORIQUE DE L'ETUDE 172
SECTION 2. METHODOLOGIE 177
SECTION 3. ANALYSES ET RESULTATS 178
SECTION 4. CONCLUSION 181
SECTION 5. LIMITES 183

CONCLUSION GENERALE **185**

SECTION 1. CONTRIBUTIONS 186
SECTION 2. LIMITES ET VOIES DE RECHERCHE 190

REFERENCES **195**

– Introduction –

La question de l'évaluation globale est au cœur des préoccupations des praticiens comme des consommateurs. Après chaque achat en ligne, chaque réservation, les clients sont sollicités pour évaluer le produit ou le service qu'ils ont acheté, leur expérience sur le site, la livraison, etc. Le développement du commerce en ligne repose en partie sur les évaluations de consommateurs qui permettent de savoir si le distributeur ou le prestataire sont dignes de confiance. Bien souvent limitées à quelques étoiles, les échelles proposées par les sites de e-commerce sont, dans la plupart des cas, assimilables à des différentiels sémantiques. Cet outil très simple oppose aux deux extrémités d'un continuum bipolaire une paire d'adjectifs antagonistes, par exemple « satisfait » à « insatisfait » ou « positif » à « négatif ». Les différentiels sémantiques soulèvent un problème largement discuté dans la littérature méthodologique tant en psychologie qu'en marketing (voir par exemple Kaplan, 1972 ; Olsen, 1999 ; Pinson, 1983 ; Thompson, Zanna et Griffin, 1995) : pour se situer sur un point du continuum, les répondants doivent effectuer le solde net entre leur niveau de positivité et leur niveau de négativité. Cette opération est problématique car elle conduit à agréger au niveau du point central des différentiels sémantiques les réponses indifférentes et les réponses ambivalentes (figure 1). La distinction entre ces deux types d'évaluations est un enjeu, dans la mesure où elles recouvrent des réalités distinctes.

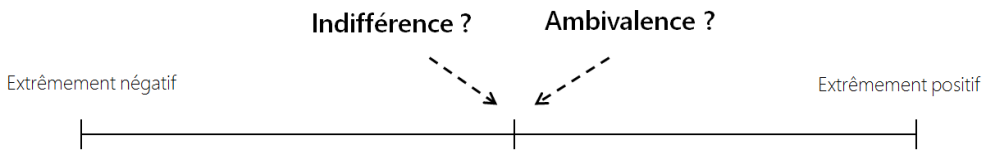


Figure 1. Réponses agrégées au niveau du point central des différentiels sémantiques

Décrivant une évaluation composée à la fois de réactions positives et de réactions négatives (Thompson, Zanna et Griffin, 1995), l'ambivalence des consommateurs est un phénomène courant dès lors que plusieurs éléments entrent en ligne de compte dans l'évaluation. Pour ne citer qu'un exemple, les enjeux de la société moderne favorisent ces situations de tiraillement : le *low cost* met l'accent sur le prix au détriment du confort, alors que l'achat de produits fabriqués dans des conditions respectueuses des salariés et de l'environnement suppose de se rendre dans des points de vente spécialisés et de payer plus cher. Entre un discours centré sur la préservation du pouvoir d'achat et la nécessité d'adopter des comportements « responsables », les consommateurs sont sans cesse amenés à mettre en balance plusieurs attributs pour former leur évaluation.

Ce phénomène pose un problème de mesure car il entre en contradiction avec la vision bipolaire des évaluations proposée par Osgood et ses collègues (1957). Selon eux, les variations sémantiques peuvent se résumer à une seule dimension évaluative, la « valence », opposant le « mauvais » au « bon ». Cette approche bipolaire ou unidimensionnelle des évaluations a donné lieu au développement des différentiels sémantiques, outil de mesure très apprécié du marketing. Sur cet outil de mesure, plus la réponse donnée est positive, moins elle est négative, et inversement. Cette approche exclut donc les évaluations composées de réactions à la fois positives et négatives.

L'approche bipolaire ou unidimensionnelle est remise en cause par de récents travaux en psychologie sur le système évaluatif des êtres humains. Ces recherches démontrent que les mécanismes d'évaluation de la positivité et de la négativité répondent à deux systèmes distincts (Cacioppo et Berntson, 1994 ; Cacioppo & al., 2011 ; Cacioppo, Gardner et Berntson, 1997). Cette conception est cohérente avec des recherches existantes en sciences de gestion. Les travaux d'Herzberg et ses collègues (1959) montrent que les facteurs de motivation et de démotivation au travail sont distincts. De même en marketing, le modèle tétraclasse de Llosa envisage séparément les attributs contribuant à la satisfaction et les attributs contribuant à l'insatisfaction dans le cadre de l'évaluation d'un service (Lichtlé,

Llosa et Plichon, 2002 ; Llosa, 1996).Ainsi, en sciences de gestion, l'idée suivant laquelle la satisfaction est un construit bidimensionnel, c'est-à-dire que les réactions positives de satisfaction et négatives d'insatisfaction sont deux concepts au moins partiellement distincts, est largement admise (Ray et Gotteland, 2005).

L'*Evaluative Space Grid* (Larsen & al., 2009), outil de mesure issu de recherches en psychologie, s'inscrit dans la droite ligne de l'approche bidimensionnelle. Il s'agit d'une mesure des opinions positives et négatives, effectuée à l'aide d'une grille à deux dimensions. Nous souhaitons montrer dans cette thèse comment transposer cet outil à la mesure des réactions globales de consommateurs. En effet, en mesurant séparément les opinions positives et les opinions négatives, cet outil devrait permettre de distinguer les quatre types de réactions évaluatives qui découlent de l'approche bidimensionnelle.

- Les évaluations globalement positives, correspondant à des réactions positives fortes et des réactions négatives faibles ou inexistantes ;
- Les évaluations globalement négatives, correspondant à des réactions négatives fortes et des réactions positives faibles ou inexistantes ;
- Les évaluations globalement indifférentes, correspondant à des réactions positives faibles et des réactions négatives faibles ;
- Les évaluations globalement ambivalentes, correspondant à des réactions positives modérées à fortes et des réactions négatives modérées à fortes.

Formulation de la problématique de recherche

Dans la droite ligne de l'approche bidimensionnelle, nous soutenons dans cette thèse que l'ambivalence constitue une partie du spectre évaluatif. L'enjeu de la thèse est donc de mesurer les évaluations des consommateurs tout en captant cette part d'ambivalence, et ce grâce à l'*Evaluative Space Grid*, outil de mesure alternatif. Cet objectif nous amène à formuler la problématique suivante :

Saisir la part d'ambivalence contenue dans les évaluations globales
en utilisant l'*Evaluative Space Grid*, permet-il d'améliorer
la prédiction des intentions de comportement des consommateurs ?

Cette problématique générale peut être déclinée en différents objectifs de recherche :

- *Objectif 1* : L'*Evaluative Space Grid* permet-elle de rendre compte des évaluations des consommateurs? La première étape de notre travail consiste à s'assurer que cet outil permet effectivement de mesurer les construits qu'il entend apprécier. Pour ce faire, nous devons établir la validité convergente de l'*Evaluative Space Grid* avec les outils de mesure traditionnels des évaluations de consommateurs. Plus précisément, cette démarche consiste à vérifier que les évaluations obtenues avec l'*Evaluative Space Grid* sont suffisamment corrélées avec celles issues des principaux outils de mesure des réactions globales et de l'ambivalence des consommateurs.

- *Objectif 2* : Comparée au différentiel sémantique, l'*Evaluative Space Grid* offre-t-elle une meilleure prédiction des intentions de comportement des consommateurs ? En appliquant l'*Evaluative Space Grid* à la mesure de la satisfaction client, nous souhaitons démontrer la supériorité psychométrique de ce nouvel outil par rapport à l'outil de mesure traditionnel des réactions de consommateur. Dans le cadre d'une approche marketing, le pouvoir prédictif de l'*Evaluative Space Grid* constitue l'enjeu majeur de la démonstration à cette étape.

- *Objectif 3* : Comment les répondants s'approprient-ils l'*Evaluative Space Grid* lorsqu'ils se trouvent en situation de répondre à un questionnaire ? Nous abordons pour finir la question du sens donné par les individus aux différentes modalités de réponses de ce nouvel outil.

Contributions attendues

Les contributions que nous souhaitons apporter sont de différentes natures.

La thèse se donne avant tout pour objectif de fournir une contribution méthodologique à la littérature en marketing en important de la psychologie un outil de mesure permettant de résoudre la question du point central des différentiels sémantiques. Plus qu'une simple transposition d'une discipline à une autre, notre travail vise à comprendre comment cet outil de mesure pourrait trouver sa place dans la conduite d'études et de recherches en marketing. En effet, si l'outil semble présenter les qualités nécessaires pour distinguer les

évaluations indifférentes des évaluations ambivalentes, nous souhaitons mettre en évidence la manière dont les répondants se l'approprient lorsqu'ils répondent à des questionnaires.

Au niveau théorique, ce travail cherche à contribuer à l'approche bidimensionnelle des réactions évaluatives en intégrant le concept d'ambivalence au spectre des évaluations globales de consommateurs. En effet, si la littérature en psychologie et en gestion reconnaît la nécessité d'envisager séparément les réactions positives et les réactions négatives, cela n'a pas été mis jusqu'ici en lien avec l'ambivalence des consommateurs. Nous souhaitons donc proposer une nouvelle conception des évaluations globales de consommateurs impliquant quatre types de réactions : les réactions positives, négatives, indifférentes et ambivalentes.

Enfin, au niveau managérial, l'introduction de l'*Evaluative Space Gr*id a pour objectif de mieux anticiper les comportements. Plus précisément, en distinguant l'indifférence et de l'ambivalence, les praticiens seront en mesure de se focaliser uniquement sur des évaluations prédictives des intentions des consommateurs.

Plan de la thèse

Nous allons tout d'abord présenter les recherches portant sur le concept d'ambivalence et sa mesure. Cette revue de la littérature sera complétée par une étude préliminaire sur les outils utilisés dans le e-commerce pour mesurer les évaluations globales (Chapitre 1). Cela nous permettra de mettre en évidence les apports potentiels de l'*Evaluative Space Grid* dans la conduite d'études consommateurs (Chapitre 2).

La suite du document sera consacrée aux cinq études empiriques permettant de répondre progressivement aux trois objectifs issus de la problématique générale de la thèse.

- Pour répondre au premier objectif de la recherche, les études 1 et 2 (Chapitre 3) seront consacrées au test de la validité convergente de ce nouvel l'outil avec les outils de mesure existants des évaluations globales et de l'ambivalence des consommateurs.

- Pour répondre au deuxième objectif de la thèse, l'étude 3 (Chapitre 4) aura pour but de démontrer empiriquement la supériorité de l'*Evaluative Space Grid* par rapport au différentiel sémantique. Plus précisément, nous appliquerons cet outil à

la mesure de la satisfaction client pour saisir la part d'ambivalence contenue dans ce construit bidimensionnel.

- Après avoir démontré que l'*Evaluative Space Grid* présente des qualités psychométriques aussi satisfaisantes que le différentiel sémantique, nous essaierons de comprendre comment les répondants s'approprient l'outil dans le cadre d'études consommateurs. Pour répondre à ce troisième objectif de la recherche, nous présenterons les résultats d'une étude qualitative (Etude 4, Chapitre 5) réalisée auprès d'un panel d'experts. Cette phase qualitative permet de clarifier le sens attribué aux différentes modalités de réponses de l'outil. Les propositions issues de cette étude seront ensuite testées dans le cadre d'une étude quantitative permettant de valider la présence d'une zone évaluative d'incertitude au niveau du point central de l'*Evaluative Space grid* (Etude 5, Chapitre 6).

Ce travail se conclut par l'exposé des principales contributions, limites et voies de recherches qui découlent de la thèse.

La figure 2 décrit le plan de la thèse (voir page 18).

Connaissez-vous cette histoire frivole
D'un certain âne illustre dans l'école?
Dans l'écurie on vint lui présenter
Pour son dîner deux mesures égales,
De même force, à pareils intervalles;
Des deux côtés l'âne se vit tenter
Également, et, dressant ses oreilles,
Juste au milieu des deux formes pareilles,
De l'équilibre accomplissant les lois,
Mourut de faim, de peur de faire un choix.

Voltaire, « La Pucelle d'Orléans », œuvre en 21 chants, chant XII, vers 16 et sq.
Œuvres complètes de Voltaire, tome XI, Paris, 1784

Chapitre 1	
Revue de la littérature	Le concept d'ambivalence La mesure de l'ambivalence
Etude préliminaire	La mesure des évaluations globales de consommateurs dans le e-commerce : pratiques et perception des internautes

↓

Chapitre 2	
Cadre conceptuel et hypothèses	Présentation de *l'Evaluative Space Grid* Résultats attendus

↓

Chapitre 3	
Etude 1	Validité convergente de *l'Evaluative Space Grid* avec les paires de continuums unipolaires
Etude 2	Validité convergente de *l'Evaluative Space Grid* avec les différentiels sémantiques, les discours des consommateurs et une mesure de l'ambivalence subjective

↓

Chapitre 4	
Etude 3	Test de *l'Evaluative Space Grid* comme outil de mesure de la satisfaction client

↓

Chapitre 5	
Etude 4	Etude qualitative après d'un échantillon d'experts : - qu'en pensent les chercheurs ? - qu'en pensent les professionnels ?

↓

Chapitre 6	
Etude 5	Place de l'incertitude dans les évaluations effectuées avec *l'Evaluative Space Grid*

Figure 2. Plan de la thèse

– Chapitre 1 –

Définition et mesure de l'ambivalence

Ce premier chapitre se donne pour objectif de synthétiser la manière dont l'ambivalence et sa mesure sont envisagées dans les recherches en psychologie et en marketing. Il se compose de deux parties. La première section vise à préciser le concept d'ambivalence. La deuxième section présente les outils de mesure de l'ambivalence, leurs limites et les implications théoriques qui en découlent.

Section 1. Qu'est-ce que l'ambivalence ?

Le terme « ambivalence » est issu du latin, « *ambi* » préfixe exprimant l'idée de « deux à la fois » et « *valens* » qui signifie « fort », « énergique », « efficace ». Le *Petit Larousse*[1] définit l'ambivalence comme une réaction pouvant conduire « à éprouver ou à manifester simultanément deux sentiments opposés à l'égard d'un même objet : amour et haine, joie et tristesse, etc. ». Si cette définition n'a rien d'académique, elle décrit bien les situations du quotidien dans lesquelles les individus sont confrontés à un déluge d'informations divergentes à propos des faits d'actualité, des objets qu'ils utilisent ou des idées qui circulent à un moment donné.

Pour approfondir la connaissance de ce concept, nous distinguerons dans cette section les contributions des recherches en psychologie (1.), les apports de la littérature en marketing

[1] *Ouvrage Collectif (2013), Le Petit Larousse Illustré, Editions Larousse, Paris*.

sur l'ambivalence des consommateurs (2.), les domaines dans lesquels ce phénomène a été étudié (3.), ses propriétés (4.) et les conséquences de l'ambivalence sur le plan des comportements (5.).

1. L'ambivalence en psychologie

Dans une perspective large, l'*American Psychology Association* (APA) définit, dans son *Dictionnaire de la Psychologie*[2], l'ambivalence comme « l'existence simultanée d'émotions et d'attitudes contradictoires, telles que le plaisir et le déplaisir ou la sympathie et l'hostilité envers la même personne, le même objet, le même événement ou la même situation ». Plus précisément, l'APA décrit *l'ambivalence d'une attitude* comme la « propension avec laquelle les réponses évaluatives associées à une attitude sont incompatibles entre elles. Si les réponses évaluatives sont uniformément positives ou négatives, l'ambivalence est faible. Si l'attitude est composée de réponses positives et négatives, l'ambivalence est élevée ».

En psychologie clinique, c'est le psychiatre suisse Eugen Bleuler (1950) qui a défini en premier ce terme dans le cadre de ses études sur la schizophrénie. Il parlait alors d'*ambivalence affective*, notamment pour qualifier les sentiments extrêmes que pouvaient ressentir certains de ses patients à l'égard de leurs parents (un amour profond mêlé à l'envie de tuer leur mère, par exemple).

C'est Scott (1966 ; 1968) qui a introduit le concept d'ambivalence en psychologie sociale. Il l'a décrite comme une propriété de l'attitude, l'inscrivant ainsi dans les travaux sur les réponses évaluatives. De ses réflexions ressort une définition de travail permettant de caractériser l'attitude ambivalente. Selon lui, plus les réactions favorables et défavorables sont fortes et de niveau similaire, plus le degré d'ambivalence est important (Scott, 1966). Thompson, Zanna, et Griffin (1995) dans leur chapitre consacré à l'ambivalence dans l'ouvrage de Krosnick et Petty sur la force de l'attitude (1995b), ont proposé une version simplifiée de cette définition en considérant que l'attitude d'un individu est ambivalente lorsqu'elle est composée à la fois de réactions positives et négatives. Ce chapitre est aujourd'hui considéré comme le texte de référence dans le domaine. Ainsi, on peut aimer le chocolat en angoissant à l'idée de grossir ; reconnaître l'intérêt du tri sélectif et jeter tous ses déchets dans une seule poubelle parce que c'est plus simple ; aimer profondément ses

[2] *Vandendos, G. R. (2007), APA Dictionary of Psychology, American Psychological Association, New York.*

parents tout en souhaitant vivement rentrer chez soi deux heures après le début d'un weekend en famille.

La littérature sur l'ambivalence aborde deux questions permettant de préciser ces premières définitions. Le premier point de discussion concerne la nature de l'ambivalence. En effet, pour certains auteurs (Kernberg, 1976 ; Meehl, 1964 ; Sincoff, 1990), il n'est pas nécessaire que les réactions tendent à la fois vers le positif et vers le négatif, pour parler d'ambivalence. Le simple fait de ressentir deux émotions différentes de même direction (un mélange d'émotions négatives comme de la colère et de la tristesse, par exemple), suffirait à caractériser une attitude d'ambivalente.

Nous avons écarté cette approche étendue du concept au profit de la définition restreinte de l'ambivalence proposée par Conner et Spark dans leur article séminal (2002). Cette dernière, qui fait aujourd'hui autorité en psychologie sociale, établit les deux critères suivants pour qualifier une attitude d'ambivalente :

- *Le critère de similarité*. Ce critère est satisfait lorsque les niveaux de positivité et négativité éprouvés à l'égard d'un objet tendent à être similaires. Dès lors que le niveau de positivité n'est plus le même que celui de négativité, l'attitude tend à se polariser vers le positif ou vers le négatif.

- *Le critère d'intensité*. Ce critère est satisfait dès lors que les réactions positives et négatives à l'égard de l'objet évalué atteignent au minimum une *intensité modérée*. En cas contraire, lorsque les réactions positives et négatives sont de faible intensité ou inexistantes, on peut considérer qu'il ne s'agit pas d'ambivalence, mais d'indifférence.

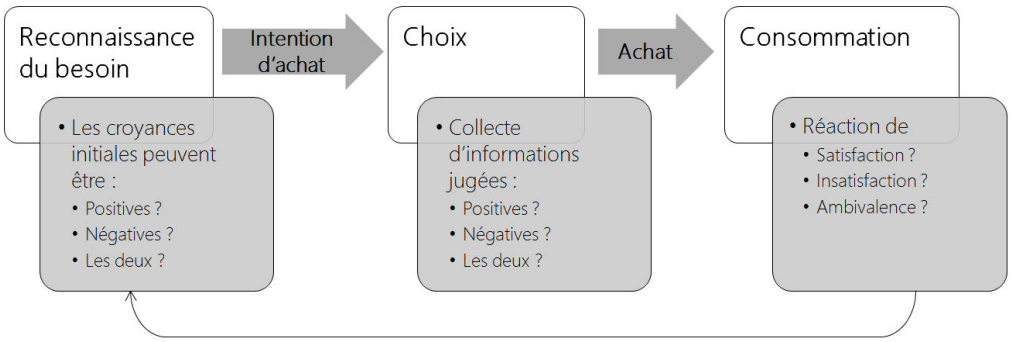
Le second point de discussion porte sur la temporalité durant laquelle sont éprouvés ces différents ressentis : cooccurrence simultanée ou séquentielle ? La réponse à cette question dépend du niveau d'analyse de l'objet d'attitude. Une analyse micro incite à se focaliser uniquement sur une partie ou un élément de l'objet évalué. Dans cette perspective, il est possible d'envisager une chaîne de réactions d'abord positives à l'égard d'une partie de l'objet, puis négatives à l'égard d'une autre partie de l'objet. Au contraire, une analyse macro invite à considérer l'objet dans sa globalité en prenant en compte plusieurs éléments dans l'évaluation. La cooccurrence peut alors être simultanée. En effet, la perception holistique augmente l'ambivalence dans la mesure où elle incite l'individu à agglomérer des

réactions (parfois contradictoires) suscitées par les différentes parties de l'objet (Olsen, 1999).

Cette thèse adopte la deuxième perspective. En effet, nous envisageons l'ambivalence au niveau des évaluations globales, agrégées. Ce niveau d'analyse nous permet de nous affranchir de la question de la temporalité des réactions.

2. L'ambivalence des consommateurs

Dans le prolongement des travaux de Thompson, Zanna et Griffin (1995), Otnes, Lowrey, Shrum (1997) ont importé le concept d'ambivalence dans le domaine de la consommation. Ils proposent de définir ce phénomène comme le fait, pour un consommateur, de « ressentir simultanément ou de manière séquentielle de multiples états émotionnels résultant de l'interaction entre des facteurs internes et des stimuli externes, dans un contexte marketing, ceci ayant des conséquences directes ou indirectes sur les attitudes et les comportements des consommateurs ». Nous proposons de simplifier cette définition pour rester davantage fidèle à celle de Thompson, Zanna et Griffin (1995) en considérant qu'**un consommateur est ambivalent lorsqu'il éprouve à la fois des réactions positives et négatives à l'égard d'un objet de consommation**.



[*Source : adapté de Darpy, 2012*]

Figure 3. La place de l'ambivalence dans le processus de consommation

L'ambivalence des consommateurs peut se manifester à tous les niveaux du processus de consommation (figure 3) : en amont de l'achat, lorsque le consommateur évalue les informations disponibles dans le but de prendre une décision, ou en aval, lorsqu'il estime son niveau de satisfaction. La question de la satisfaction client semble particulièrement

propice à l'étude de l'ambivalence dans la mesure où il est fréquent d'éprouver des réactions positives ou négatives à l'égard d'un produit ou d'un service consommé (Olsen, Wilcox et Olsson, 2005 ; Priester, Petty et Park, 2007 ; Sparks & al., 2001). C'est pourquoi nous nous intéresserons plus particulièrement dans cette thèse aux réactions ambivalentes postérieures à l'achat et/ou la consommation.

3. Les domaines d'étude de l'ambivalence

Les recherches sur l'ambivalence ont mis en évidence des réactions de direction opposée à l'égard d'un même objet d'attitude dans des domaines divers et variés.

Les « questions de société » constituent le domaine de prédilection de ces travaux. Le don de sang (Armitage, 2003 ; Priester et Petty, 2001), la peine de mort (Newby-Clark, McGregor et Zanna, 2002 ; Thompson, Zanna et Griffin, 1995), l'euthanasie (Schneider & al., 2013), l'avortement (Newby-Clark, McGregor et Zanna, 2002), la loi et les impôts (Clark, Wegener et Fabrigar, 2008 ; Schneider & al., 2013 ; van Harreveld, van der Pligt et de Liver, 2009), l'énergie nucléaire (Clark, Wegener et Fabrigar, 2008), l'utilisation du préservatif (Priester et Petty, 2001), les attitudes vis-à-vis des immigrés et des groupes minoritaires (Bell et Esses, 2002 ; Maio, Bell et Esses, 1996) ou encore l'infidélité conjugale (Kachadourian, Fincham et Davila, 2005) ont fait l'objet de publications sur l'ambivalence.

Dans une perspective marketing, les enjeux de la société moderne sont de nature à favoriser l'ambivalence des consommateurs (van Harreveld & al., 2012 ; van Harreveld, van der Pligt et de Liver, 2009). En effet, les individus doivent « jongler » entre un discours marketing exaltant le plaisir de la consommation et les injonctions consuméristes qui les incitent à adopter des comportements dits « responsables » : protéger l'environnement (trier leurs déchets, faire des économies d'énergie, etc.), préserver leur santé (ne pas manger trop gras ni trop sucré, faire du sport, etc.) et celle des autres (ne pas fumer, respecter les limitations de vitesse, etc.). Ainsi, les recherches menées jusqu'ici sur l'ambivalence des consommateurs ont exploré l'ambivalence manifestée à l'égard de certains comportements de consommation : les régimes alimentaires (Armitage et Conner, 2000 ; Conner et Sparks, 2002 ; Costarelli et Colloca, 2004 ; Priester et Petty, 2001), le tabagisme (van Harreveld & al., 2004 ; Zhao et Cai, 2008), la malbouffe (Priester et Petty, 2001), la consommation d'alcool (Armitage, 2003), les décisions d'achat effectuées lors du mariage (Otnes, Lowrey et Shrum, 1997) ou les comportements pro-environnementaux

(Costarelli et Colloca, 2004). On peut citer également des recherches sur des évaluations ambivalentes à l'égard des tabous publicitaires (Zaaraoui-Sabri, 2007), des supermarchés (Clarkson, Tormala et Rucker, 2008), mais aussi de produits comme le shampooing (Jonas, Diehl et Brömer, 1997), l'iPod (Priester, Petty et Park, 2007), le magnétoscope (Priester, Petty et Park, 2007), les aliments génétiquement modifiés (Nordgren, van Harreveld et van der Pligt, 2006 ; van Harreveld & al., 2004), la viande, le chocolat (Sparks & al., 2001), ou encore les produits de la mer (Olsen, Wilcox et Olsson, 2005).

A l'exception du travail de Clarkson, Tormala et Rucker (2008) sur les supermarchés, nous constatons que l'ambivalence à l'égard des services est un domaine d'application presque ignoré de ces travaux. C'est l'une des raisons pour lesquelles nous avons décidé de nous pencher sur ce secteur.

4. Les propriétés associées à l'ambivalence

L'étude de l'ambivalence s'inscrivant en psychologie dans le champ des travaux sur l'attitude, les chercheurs se sont attachés à définir ses propriétés comme cela se fait traditionnellement dans l'étude des réponses évaluatives. Cette discussion s'organise autour de la question de la force de l'attitude ambivalente. Selon Krosnick et Petty (1995a), la force d'une attitude peut être estimée à la lumière de quatre éléments : sa persistance dans le temps, sa résistance aux tentatives de persuasion, son influence sur le traitement de l'information et son effet sur les comportements. La question des comportements étant particulièrement importante pour le marketing, elle fait l'objet d'un développement séparé dans le point suivant.

Tout d'abord, les auteurs notent que les réactions ambivalentes sont moins stables dans le temps que les réactions polarisées. L'une des explications résiderait dans la difficulté d'accès aux réactions positives et négatives composant l'ambivalence. Cette difficulté d'accès transparait dans le temps de réponse des sujets ambivalents (Bargh & al., 1992). Dans cette perspective, les évaluations ambivalentes ne seraient pas profondément ancrées dans l'esprit des individus et pourraient évoluer. Plus précisément, selon Katz (1944), les attitudes ambivalentes seraient ainsi amenées à passer du positif au négatif en fonction de l'accessibilité relative des réactions plaisantes ou déplaisantes au fil du temps ou selon le contexte.

L'étude de la résistance aux tentatives de persuasion des individus ambivalents est en lien avec la manière dont ils traitent les informations susceptibles de les faire changer d'avis. La littérature n'est pas homogène sur cette question. Jonas, Diehl et Brömer (1997) ont montré que les évaluations ambivalentes favorisent un traitement systématique de l'information. En effet, l'ambivalence génèrerait un sentiment d'inconfort (Nordgren, van Harreveld et van der Pligt, 2006) qui motiverait les individus à résoudre le conflit attitudinal auquel ils se heurtent de manière à rétablir une situation de cohérence interne (Judd et Brauer, 1995 ; McGregor, Newby-Clark et Zanna, 1999 ; Nordgren, van Harreveld et van der Pligt, 2006). Au contraire, en se penchant plus précisément sur la question de la résistance à la persuasion, Zemborain et Johar, (2007) ont montré que les personnes qui ont une attitude ambivalente à l'égard d'une question sont plus influençables que les personnes ayant une attitude polarisée. Plus précisément, dans une étude portant sur l'appréciation d'un candidat politique, ils ont démontré que contrairement aux autres, les participants ambivalents ne faisaient pas de distinction entre une source fiable et une source moins fiable avant d'accepter un message.

5. Les conséquences comportementales de l'ambivalence

La question de l'impact comportemental de l'ambivalence est centrale dans les recherches sur le sujet, en particulier dans une perspective marketing. La littérature a abordé ce sujet de deux manières : soit l'ambivalence est considérée comme un antécédent des intentions de comportement, soit elle est traitée comme un modérateur de la relation entre attitude et intentions de comportement. Les résultats des études traitant du sujet diffèrent selon l'angle d'analyse adopté.

La première approche consiste à envisager l'ambivalence comme un antécédent direct des intentions de comportements. Elle compte peu de travaux et ceux-ci n'ont pas souvent abouti à des résultats significatifs (Olsen, Wilcox et Olsson, 2005). Dans la recherche de Sparks et ses collègues (2001), la relation entre l'ambivalence à l'égard du chocolat et l'intention d'en consommer n'est pas significative, alors que les auteurs ont trouvé une relation négative significative entre l'ambivalence et l'intention de consommer de la viande. Dans un tout autre registre, Costarelli et Colloca (2004) ont montré que l'ambivalence à l'égard des comportements pro-environnementaux est corrélée négativement avec l'intention de les adopter. Olsen et ses collègues (2005) se sont quant à eux intéressés à l'impact de l'ambivalence sur l'intention de fidélité. Ils ont mis en évidence une relation

négative significative entre l'ambivalence à l'égard des produits de la mer et l'intention d'en racheter.

Une autre série d'étude, plus abondante, montre que l'ambivalence diminue la capacité de l'attitude à prédire les comportements. On retrouve ici l'idée de la faiblesse des attitudes ambivalentes (Armitage et Conner, 2000 ; Conner et Sparks, 2002 ; Thompson, Zanna et Griffin, 1995 ; Zemborain et Johar, 2007), attitudes qui ne seraient pas profondément ancrées dans l'esprit des individus et qui pourraient être amenées à évoluer. Krosnick et Petty (1995a) considèrent ainsi que l'ambivalence ne peut être envisagée comme une variable pertinente pour prédire les comportements. Au contraire, Jonas, Diehl et Brömer (1997), ainsi que Sengupta et Johar (2002) affirment que l'ambivalence améliore la relation attitude-comportement. En mobilisant les modèles duels de traitement de l'information (Brewer, 1988 ; Petty et Cacioppo, 1986), ces auteurs ont montré que les individus confrontés à un stimulus ambivalent tendent à traiter les informations de manière systématique. Or, les attitudes formées via ce processus sont plus prédictives des intentions de comportement (Petty, Priester et Wegener, 1994).

Il semble, pour résumer, que l'ambivalence puisse avoir deux types d'effets sur les intentions de comportements des individus. De rares études témoignent d'un effet négatif de l'ambivalence sur les intentions de comportement. L'ambivalence pousserait alors les individus à la prudence, ce qui se traduirait par des comportements d'évitement. Une autre série de recherches attribue à l'ambivalence un effet sur la force du lien attitude-intention de comportement. Dans cette perspective, les individus tendraient à réagir de manière proactive à la tension générée par les évaluations ambivalentes en se livrant à un traitement plus systématique de l'information.

Section 2. Comment mesurer l'ambivalence ?

L'intérêt croissant des chercheurs en sciences humaines et sociales pour l'ambivalence les a conduits à soulever à plusieurs reprises la question de sa mesure (voir par exemple : Olsen, 1999 ; Olsen, Wilcox et Olsson, 2005 ; Priester et Petty, 1996 ; Thompson, Zanna et Griffin, 1995).

Comme nous l'avons expliqué en introduction, notre objectif est de saisir la part d'ambivalence contenue dans les évaluations globales de consommateurs. Nous évacuons donc immédiatement les mesures dites « subjectives » de l'ambivalence. Ces mesures ont

pour but de capturer un ressenti d'ambivalence (Conner et Sparks, 2002 ; Lavine, Borgida et Sullivan, 2000 ; Priester et Petty, 1996). Il s'agit de demander directement aux sujets dans quelle mesure ils nourrissent des opinions de direction opposée à l'égard de l'objet évalué. Plusieurs échelles multi-items ont été développées à cette fin. Conner et Sparks (2002) ont par exemple proposé les items suivants : « j'ai des opinions contradictoires au sujet de… », « j'ai un avis partagé vis-à-vis de… », « je suis indécis au sujet de… », « j'ai des opinions positives et négatives au sujet de… ». Les mesures subjectives reposent sur la perception du répondant. Elles supposent donc qu'il soit conscient de son ambivalence (Jonas, Brömer et Diehl, 2000). Le terme employé dans la littérature pour qualifier le produit de ces mesures est celui de *sentiment d'ambivalence*, en opposition à une mesure de l'*ambivalence attitudinale*. Cette approche étant très éloignée de celle que nous développons dans cette thèse, nous ne nous étendrons pas davantage sur ce type de mesures. Elles pourront toutefois s'avérer utiles ultérieurement dans une perspective de comparaison avec les données issues de l'outil que nous souhaitons tester.

Plutôt que d'étudier les mesures « subjectives » nous allons nous pencher sur les mesures dites « objectives » de l'ambivalence qui consistent à capturer séparément les réactions positives et les réactions négatives des individus (Conner et Sparks, 2002). Leur développement est intéressant parce qu'il est intrinsèquement lié à un débat sur la nature de l'attitude. Or, cette discussion peut être rapprochée de la question de la mesure des évaluations globales de consommateurs. Nous allons suivre le développement chronologique de la littérature en présentant d'abord dans cette section les discussions portant sur la bipolarité des affects et l'outil associé à cette approche, le différentiel sémantique (1.). Nous verrons ensuite pourquoi cet outil ne permet pas de mesurer l'ambivalence (2) et nous présenterons les solutions de mesures proposées dans la littérature pour répondre à ce problème (3). Enfin, nous conclurons sur l'apport théorique d'une telle approche (4).

1. Le différentiel sémantique, outil de mesure bipolaire

La technique du différentiel sémantique trouve ses origines dans les recherches sur la synesthésie. La synesthésie est un phénomène neurologique par lequel plusieurs sens sont associés. Par exemple, dans un type de synesthésie connu sous le nom de synesthésie « graphèmes-couleurs », les lettres de l'alphabet ou les nombres peuvent être associés à certaines couleurs. Si la véritable synesthésie neuronale est involontaire et concerne peu de

personnes (entre 4% et 13% de la population selon les études), ce champ de recherche a mis en évidence un certain nombre de « métaphores synesthésiques ». Il s'agit d'associations sensorielles largement partagées par les individus, au moins dans une culture donnée. Par exemple, pour les Occidentaux ce qui est chaud est rouge, les musiques lentes et mélancoliques sont visualisées par des formes sombres, lourdes et massives, l'espoir est blanc et il y a une association entre « proche » et « rapide ».

Osgood (1952) en a conclu que l'imagerie mentale au cœur de ces synesthésies permet de mettre en évidence des relations sémantiques entre ce qu'il appelle des « continuums ». Ces continuums décrivent des variations sémantiques opposant deux termes polaires, le plus souvent des paires d'adjectifs antonymes. Plutôt que d'étudier des relations entre des adjectifs, il propose d'utiliser ces « continuums » ou « différentiels sémantiques » pour établir des relations entre des adjectifs et des concepts. Avec l'aide de Stagner (1946), il a collecté entre avril 1940 et mars 1942 des données sur le sens associé aux concepts de « pacifiste », « russe », « dictateur » et « neutralité ». Les analyses menées leur ont permis d'identifier plusieurs « grappes » de signification. En effet, les paires « juste-injuste », « haut-bas », « gentil-cruel », « précieux-sans valeur » et « chrétien-antichrétien » étaient corrélées à hauteur de 0,9. Ils ont qualifié cette « grappe » de jugements allant globalement du « bon » au « mauvais », de dimension « évaluative ». Cette dimension sera plus tard désignée sous le terme de « valence » des évaluations. L'existence de cette dimension évaluative a été confirmée par Osgood, Suci et Tanenbaum (1957) qui ont également mis en évidence deux autres dimensions à partir d'analyses factorielles : le facteur d'activité (« calme-excité ») et le facteur de dominance. Ce dernier a progressivement été abandonné car il ne ressortait pas toujours comme une dimension clairement séparée des deux autres facteurs. Quoi qu'il en soit, c'est la dimension évaluative « bon-mauvais » qui restitue la plus grande part de la variance.

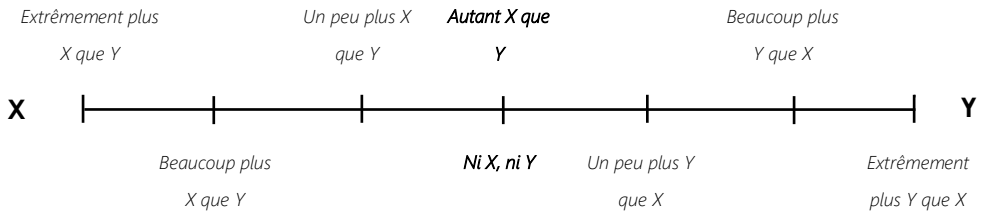
Selon Osgood, l'espace sémantique est donc globalement bipolaire et il existe une multitude d'adjectifs antagonistes réciproques. Quelle que soit la manière dont le sens varie sur les continuums testés dans ses études (Osgood, 1952 ; Stragner et Osgood, 1946), les variations sémantiques peuvent se résumer à une seule dimension. Il y a donc une équivalence fonctionnelle entre une multitude de paires d'adjectifs, c'est ce qui rend possible le développement d'une mesure quantitative du sens. Si au contraire tous les termes que nous utilisons étaient associés à une signification unique et indépendante de celle des autres mots, tout espoir de mesure serait anéanti (Osgood, 1952).

Parce qu'ils sont faciles à construire, à administrer et à analyser, les différentiels sémantiques sont utilisés massivement en marketing (Pinson, 1983). Les paires d'adjectifs les plus souvent retenues pour mesurer les réactions de consommateurs opposent le positif au négatif, le plaisant au déplaisant (dimension affective de l'attitude), le bénéfique au nuisible (dimension cognitive de l'attitude), le favorable au défavorable. Les échelons intermédiaires sont indiqués par des nombres plutôt que par des étiquettes sémantiques, afin de ne pas altérer la réponse qui doit refléter le degré d'adhésion pour l'un ou l'autre des adjectifs antonymes de l'échelle (Jolibert et Jourdan, 2005). Lorsqu'un individu effectue une évaluation sur un différentiel sémantique, il indique donc, d'une part, la direction de son association en choisissant de tendre vers l'un ou l'autre des adjectifs situés aux extrémités du continuum. Il informe d'autre part, de l'intensité de son association entre l'objet évalué et cette paire d'adjectifs en choisissant un point plus ou moins proche des extrémités du continuum.

2. Les limites du différentiel sémantique

Comme nous l'avons expliqué, la technique du différentiel sémantique se fonde sur une conception bipolaire de l'attitude. Pour que cette méthode soit valide, les paires d'adjectifs antagonistes opposées aux extrémités de l'échelle doivent faire l'objet d'une opposition bipolaire puisque les répondants ne peuvent sélectionner qu'un seul point sur le continuum. Cette exigence de bipolarité ne permet pas de capturer des évaluations allant dans les deux directions du continuum (Kaplan, 1972).

Une lecture attentive d'Osgood et de ses collègues (1957), montre qu'ils avaient anticipé cette critique en proposant une définition duale du point central des différentiels sémantiques. En effet, dans leurs travaux ce point, est tantôt qualifié de « ni l'un ni l'autre », tantôt de « les deux à la fois ». Cela entraine, selon Kaplan (1972), une modification fondamentale dans l'interprétation des instructions du différentiel sémantique. Le répondant ne doit plus se positionner sur un continuum bipolaire allant d'un adjectif X à son contraire Y, mais effectuer le solde net (X-Y) entre les qualités de l'objet évalué. Il s'agit donc de savoir si le répondant penche plus vers X ou vers Y pour qualifier le concept (figure 4).



[*Source : Kaplan, 1972*]

Figure 4. Interprétation des instructions du différentiel sémantique selon Kaplan (1972)

Cette redéfinition des instructions du différentiel pose néanmoins deux problèmes. Tout d'abord, la précaution méthodologique prise par Osgood et ses collègues (1957), s'est faite au détriment de la clarté d'interprétation des évaluations situées au niveau du point central des différentiels sémantiques. En effet cette définition duale du point central entraine une fusion entre deux types d'évaluations pourtant différentes :

- les évaluations *indifférentes* : il s'agit d'évaluations pour lesquelles le répondant n'arrive pas à se prononcer, soit parce qu'il n'est pas intéressé par l'objet évalué, soit parce qu'il se sent incompétent pour l'apprécier. Il choisit alors le point central de l'échelle parce que son attitude n'est ni positive, ni négative ;

- *les évaluations ambivalentes* : il s'agit d'évaluations pour lesquelles le répondant n'arrive pas à se prononcer parce qu'il éprouve à la fois des réactions positives et négatives.

La question de l'interprétation de cette zone de flou a fait couler beaucoup d'encre (Dubois et Burns, 1975 ; Kaplan, 1972 ; Oetting, 1967 ; Pinson, 1983). Pour tenter de clarifier cette ambiguïté, Klopfler et Madden (1980) ont fait varier dans une étude expérimentale la définition du point médian. Selon qu'il était défini soit comme une réponse ambivalente (« parfois vous penchez d'un côté, parfois de l'autre »), soit comme une réponse incertaine (« vous n'arrivez pas à identifier vos sentiments »), soit comme une réponse indifférente (« cela vous est égal »), les résultats montrent que pour certains objets, la fréquence de choix du point médian dépend de sa qualification. En effet, dans cette étude, c'est lorsqu'il était défini comme « ambivalent » qu'il a été le plus fréquemment choisi et c'est lorsqu'il était défini comme « indifférent » qu'il a été le moins prisé. Cela illustre clairement comment le point médian des différentiels sémantiques peut agréger à tort des évaluations recouvrant des réalités différentes.

La définition duale du point central des différentiels sémantiques pose un second problème. Même si la précaution méthodologique prise par Osgood et ses collègues (1957) permet d'accueillir au niveau du point central les réactions ambivalentes, celles-ci sont nécessairement cantonnées à un seul point sur le continuum. Il est donc impossible de distinguer d'éventuels degrés d'ambivalence. Or, selon Scott (1966), plus les réactions positives et négatives éprouvées sont de niveau similaire et d'intensité élevée, plus l'ambivalence est forte.

3. Les solutions proposées dans la littérature

Globalement, la littérature propose, pour résoudre le problème de point central des différentiels sémantique, de mesurer séparément les réactions positives et négatives en utilisant deux continuums unipolaires. Nous allons d'abord décrire cette technique (3.1.), avant d'expliquer comment les données collectées avec ce type de dispositif permettent de calculer un score d'ambivalence (3.2.).

3.1. Les paires de continuums unipolaires

Les instructions d'Osgood et ses collègues (1957) créent donc une zone de flou au niveau du point central des différentiels sémantiques. L'élimination du point central n'est pas considérée comme une solution valable dans la mesure où il est important d'offrir aux répondants la possibilité de ne pas pencher d'un côté ou de l'autre du continuum. Par ailleurs, on peut penser que les réponses médianes seraient reportées soit légèrement à gauche, soit légèrement à droite du point central en cas de suppression de ce dernier. Ce choix forcé risquerait de créer du « bruit » dans les données et ne permet en rien de résoudre la question de la mesure de l'ambivalence (Leygue, 2009).

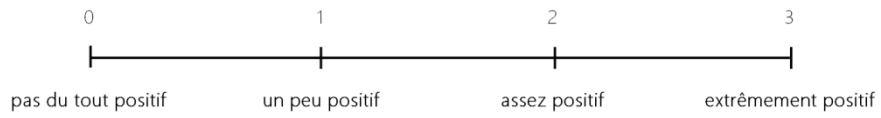
La solution généralement proposée dans la littérature consiste à mesurer l'ambivalence à l'aide de deux continuums unipolaires distincts: un continuum du plaisir ou de la positivité (allant de « pas du tout positif » à « extrêmement positif ») et un continuum du déplaisir ou de la négativité (allant de « pas du tout négatif » à « extrêmement négatif ») (Kaplan, 1972 ; Russell et Carroll, 1999). On parle alors de mesure « objective » de l'ambivalence, en opposition aux mesures dites « subjectives ». Cette technique est considérée comme « objective » en raison du caractère indirect du processus de mesure, consistant à essayer de capturer les réactions positives et négatives constituant l'ambivalence séparément. Si plusieurs méthodes existent, elles visent toutes à obtenir deux « notes » séparées, l'une

pour les aspects positifs et l'autre pour les aspects négatifs de l'objet évalué. Ces deux notes sont ensuite combinées dans une formule visant à calculer un score d'ambivalence.

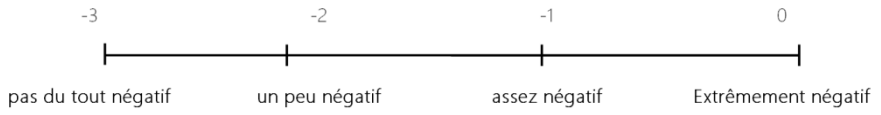
Cela suppose toutefois que les répondants acceptent d'évaluer séparément leur niveau de positivité et de négativité. Ces deux concepts étant communément considérés comme deux antagonistes réciproques, les individus pourraient être tentés de répondre de manière symétrique aux deux items. Cela se traduirait par un effet de contamination entre les deux questions : les répondants pourraient avoir tendance à ajuster la réponse apportée au deuxième continuum de manière à ce qu'elle soit cohérente avec la réponse fournie pour le premier continuum (Mucchielli, 1999). C'est pour cela que Kaplan (1972) a proposé de marquer formellement la division du différentiel sémantique en deux parties, pour former deux échelles unipolaires. Il a qualifié cette méthode de technique des composants (figure 5). De manière générique, la littérature parle de la méthode des deux continuums unipolaires.

La consigne suivante introduit chaque continuum :

« *Considérant uniquement les aspects positifs d'un concept et ignorant les aspects négatifs, veuillez évaluer dans quelle mesure ces aspects positifs sont positifs sur cette échelle en 4 points* ».



« *Considérant uniquement les aspects négatifs d'un concept et ignorant les aspects positifs, veuillez évaluer dans quelle mesure ces aspects négatifs sont négatifs sur cette échelle en 4 points* ».



[*Source : Kaplan, 1972*]

Figure 5. La technique des composants ou des deux continuums unipolaires

Russell et Carroll (1999) craignent que l'utilisation de ces deux continuums unipolaires suscite artificiellement des réponses ambivalentes par leur simple présence. Pour pallier cette limite, ils prônent l'usage d'échelles dichotomiques-puis-unipolaires. D'abord utilisée par Ekman et ses collègues (1987), puis par Reisenzein (1995) dans le cadre de leurs travaux

sur l'expression faciale des émotions, cette technique est constituée de deux étapes. Dans un premier temps on s'intéresse à l'existence d'une réaction positive avec un item du type « Etes-vous heureux ? » (« oui » ou « non »). Seuls les individus ayant répondu « oui » à la première question sont ensuite invités à indiquer l'intensité de cette réaction sur une échelle de Likert en 6 points (« Dans quelle mesure êtes-vous heureux ? » (allant de « pas du tout » à « extrêmement »). La même procédure en deux étapes est déployée pour mesurer la tristesse.

Ces techniques permettent de résoudre les deux problèmes associés aux différentiels sémantiques. En combinant les notes de positivité et de négativité, on peut, d'une part, distinguer l'indifférence (ni positif, ni négatif) de l'ambivalence (à la fois positif et négatif) et, d'autre part, distinguer différents degrés d'ambivalence (« modérément ; modérément » ou « fort ; fort »). Le paragraphe qui suit est consacré à la discussion des différentes méthodes permettant de calculer le score d'ambivalence.

3.2. Le calcul du score d'ambivalence

Plusieurs formules combinant la note de positivité (appelée « P) et la note de négativité (appelée « N ») ont été proposées pour établir le degré d'ambivalence des répondants. Les forces et les faiblesses de chacune d'entre elles ont été discutées par Thompson et ses collègues (1995) [3].

La formule de Kaplan (1972) consiste à soustraire la « polarisation » de « l'affect total ». L'affect total c'est la somme de P + N. La polarisation correspond à la valeur absolue de la différence entre P et N. Un score élevé indique que la réponse est polarisée vers le positif ou vers le négatif. Dans cette perspective, l'ambivalence correspond à la zone de chevauchement entre le continuum de la positivité et le continuum de la négativité. Elle prend des scores allant de 2 à 8, présentés dans le tableau 1.

En observant attentivement cet index, on s'aperçoit que l'ambivalence calculée correspond au double de la plus faible des deux notes (P ou N). Ce n'est pas satisfaisant dans la mesure où la note la plus faible fige le score d'ambivalence, quelle que soit la note la plus élevée. Or, on peut penser que le degré d'ambivalence d'un individu n'est pas le même lorsqu'il indique (a) une note de 2 sur les deux continuums, ou bien (b) lorsqu'il indique une note de

[3] *Dans l'article de Thomson et ses collègues (1995), les notes de positivité et de négativité étaient exprimées en valeur absolue. Pour permettre la comparaison entre les différentes formules, dans le cadre de cette discussion ces notes prennent systématiquement des valeurs allant de 1 à 4.*

2 sur le continuum de la positivité et de 4 sur celui de la négativité. Si la situation (a) décrit bien une réaction ambivalente, la situation (b) correspond à une réponse de direction négative.

$$Ambivalence = (P+N) - |P-N|$$

Continuum négatif				
4	2	4	6	8
3	2	4	6	6
2	2	4	4	4
1	2	2	2	2
	1	2	3	4

Continuum positif

[*Source : Thompson et al. 1995*]

Tableau 1. Scores d'ambivalence selon la formule de Kaplan (1972)

Katz et Hass (1988) se sont également penchés sur cette question du calcul du score d'ambivalence dans le cadre d'une recherche sur les attitudes à l'égard de la minorité noire. Ils ont proposé la formule ci-dessous prenant des valeurs allant de 1 à 16. Malheureusement, cette formule ne permet pas de résoudre les problèmes soulevés par la proposition de Kaplan. Au contraire, les réactions polarisées obtiennent des scores plus élevés que certaines réactions ambivalentes, ce qui est contraire aux exigences d'un tel index. En effet, un individu qui donnerait une évaluation de (1 ; 1) se verrait attribuer un score d'ambivalence inférieur à celui qui donnerait une évaluation polarisée de (1 ; 4) comme l'indique le tableau 2.

$$Ambivalence = P \times N$$

Continuum négatif				
4	4	8	12	16
3	3	6	9	12
2	2	4	6	8
1	1	2	3	4
	1	2	3	4

Continuum positif

[*Source : Thompson et al. 1995*]

Tableau 2. Scores d'ambivalence selon la formule de Katz et Hass (1988)

Jamieson (1988) a quant à lui proposé, de calculer un ratio entre les deux notes. Pour donner du poids à l'intensité des notes, il a suggéré de mettre au carré la note la plus faible. Son index d'ambivalence prend des notes allant de 0,25 à 4 (tableau 3).

$$Ambivalence = (note\ la\ plus\ faible)^2 / note\ la\ plus\ élevée$$

Continuum négatif				
4	0,25	1	2,25	4
3	0,33	1,33	3	2,25
2	0,5	2	1,33	1
1	1	0,50	0,33	0,25
	1	2	3	4

Continuum positif

[*Source : Thompson et al. 1995*]

Tableau 3. Scores d'ambivalence selon la formule de Jamieson (1988)

Si cette formule permet de dépasser les limites des deux précédentes, elle ne fait pas l'objet d'une réelle justification théorique. En fait, Jamieson s'est contenté de comparer les scores obtenus avec des différentiels sémantiques pour établir cet index. On peut se demander par exemple pourquoi il a proposé de mettre au carré la note la plus faible et pas au cube.

Après avoir passé en revue les différentes équations existant dans la littérature, Thomson et ses collègues (1995) fournissent une formule qui est à la fois justifiable théoriquement et conforme à la définition du concept d'ambivalence. En effet, leur proposition permet de transférer directement les conditions définissant l'ambivalence dans une équation. Pour rappel, la première condition, ou critère de similarité, est que la note de positivité et la note de négativité soient de même intensité (en cas contraire l'évaluation est polarisée vers le positif ou le négatif). La seconde condition, ou critère d'intensité, est que les deux notes soient au minimum d'intensité modérée (en cas contraire, il s'agit d'indifférence) (Conner et Sparks, 2002 ; Scott, 1966). La formule qu'ils ont proposée s'énonce comme suit :

$$Ambivalence = \overbrace{(P+N)/2}^{Intensité} - \overbrace{|P-N|}^{Similarité}$$

L'intensité est obtenue en effectuant la moyenne entre la note de positivité et la note de négativité, tandis que le calcul de la similarité correspond à la valeur absolue de la

différence entre la note P et N. La similarité est élevée quand l'intensité des deux notes est identique. Elle est faible lorsque l'intensité de la positivité et l'intensité de la négativité atteignent la différence maximale, c'est-à-dire quand l'évaluation est polarisée. L'ambivalence calculée avec cette formule prend des scores allant de -0,5 à 4.

Continuum négatif				
4	-0,5	1	2,5	4
3	0	1,5	3	2,5
2	0,5	2	1,5	1
1	1	0,5	0	-0,5
	1	2	3	4

Continuum positif

[*Source : Thompson et al. 1995*]

Tableau 4. Scores d'ambivalence selon la formule de Thompson et al. (1995)

Les scores ainsi obtenus sont très proches de ceux calculés avec l'équation de Jamieson (1988), toutefois c'est la seule formule qui permette d'estimer séparément le poids relatif de l'intensité et de la similarité des deux notes (tableau 4).

Thompson et ses collègues (1995) ont démontré la supériorité de leur équation dans le cadre d'une étude menée auprès de 146 personnes et portant sur plusieurs sujets de société. Deux mesures ont été employées : une mesure subjective du sentiment d'ambivalence effectuée à l'aide de trois items et une mesure objective effectuée avec deux continuums unipolaires de la positivité et de la négativité. Les données collectées avec les continuums ont été utilisées pour calculer les scores d'ambivalence des répondants suivant les différentes formules. Ces notes sont comparées entre elles dans la matrice de corrélation reproduite dans le tableau 5. Les résultats indiquent que le score moyen d'ambivalence varie selon la méthode utilisée. La formule de Thompson et ses collègues (1995) est celle qui aboutit au score le plus faible. Par ailleurs, c'est cette équation qui fournit les scores les plus corrélés avec la mesure subjective du ressenti d'ambivalence. C'est pourquoi c'est aujourd'hui celle qui fait autorité.

	Kaplan	Katz & Hass	Jamieson	Thompson	Mesure subjective
Kaplan	*0,68*				
Katz & Hass	0,71	*0,66*			
Jamieson	0,75	0,67	*0,61*		
Thompson	0,77	0,66	0,98	*0,58*	
Mesure subjective	0,32	0,21	0,37	*0,40*	*0,88*
Moyenne [écart type]	3,66 [0,83]	5,71 [1,52]	1,37 [0,48]	1,31 [0,58]	3,78 [1,54]

Alphas de Cronbach en gras italique. Ecart type entre parenthèses.

[*Source : Thompson et al. 1995*]

Tableau 5. Corrélations entre les différentes mesures de l'ambivalence

Cette section souligne l'intérêt d'une appréciation distincte de la positivité et la négativité d'un stimulus pour capturer les évaluations ambivalentes. Un tel constat invite à questionner l'hypothèse d'unidimensionnalité associée aux différentiels sémantiques (Osgood, 1952 ; Osgood, Suci et Tannenbaum, 1957). En effet, si certaines évaluations sont composées à la fois de réactions positives et de réactions négatives, cela signifie qu'on ne peut pas opposer de manière bipolaire le positif au négatif.

4. L'approche bidimensionnelle

La limite méthodologique soulevée par les différentiels sémantiques trouve une réponse théorique dans l'approche bidimensionnelle de l'attitude. Elle a été initiée en sciences de gestion par les travaux d'Herzberg et ses collègues (1959) sur la motivation des salariés au travail. Ces recherches ont montré que les facteurs de satisfaction et d'insatisfaction sont distincts. Plus précisément, les facteurs dits « d'hygiène », tels que les conditions de travail, les relations au sein de l'équipe et la rémunération, génèrent uniquement de l'insatisfaction lorsqu'ils n'atteignent pas les attentes des salariés, mais ils ne suscitent pas de satisfaction particulière dans le cas contraire. Seuls les facteurs dits « internes » sont source de satisfaction. Cette satisfaction est attribuable au contenu des tâches, aux éventuelles promotions obtenues, ainsi qu'à l'indépendance et l'autonomie au travail. Le même constat a été établi en marketing. En effet, Llosa et ses collègues (Lichtlé, Llosa et Plichon, 2002 ; Llosa, 1996) ont mis en évidence une asymétrie entre les éléments contribuant à la satisfaction et les éléments contribuant à l'insatisfaction dans le cadre de l'évaluation d'un service. Les attributs de la satisfaction, comme le petit bonbon au moment de l'addition au restaurant, enchantent les clients, alors qu'ils n'auraient probablement pas été insatisfaits en

l'absence de ce petit « plus ». Les attributs de l'insatisfaction comme la propreté du restaurant, entrainent en revanche une évaluation très négative en cas d'échec (si le restaurant est sale), alors qu'ils n'auraient pas généré de satisfaction particulière en cas de succès (si le restaurent avait été propre). En sciences de gestion, l'idée que la satisfaction est un construit bidimensionnel, c'est-à-dire que la satisfaction et l'insatisfaction sont deux concepts au moins partiellement distincts, est largement admise (Ray et Gotteland, 2005).

Cette idée a été confirmée par les travaux de Cacioppo et ses collègues (Cacioppo et Berntson, 1994 ; Cacioppo & al., 2011 ; Cacioppo, Gardner et Berntson, 1997) sur la modélisation de l'espace évaluatif (*Evaluative Space Model*). Ces recherches ont montré que l'évaluation des aspects plaisants et l'évaluation des aspects déplaisants ne met pas en jeu les mêmes processus : il existe une distinction entre le système d'évaluation de la positivité et le système d'évaluation de la négativité. Ces deux processus étant distincts, ils n'ont pas un fonctionnement parfaitement réciproque. Ainsi, ce sont deux zones différentes du cerveau qui s'activent en réaction à des stimuli plaisants ou déplaisants. Ce processus peut plonger les individus dans un état émotionnel composé de tristesse et de joie (Cacioppo, Gardner et Berntson, 1997 ; Larsen, McGraw et Cacioppo, 2001 ; Thompson, Zanna et Griffin, 1995). Le film « *La vie est belle* » de Roberto Benigni a été utilisé dans plusieurs études pour susciter de telles réponses (Larsen et McGraw, 2011).

Ces travaux confèrent une assise théorique aux phénomènes d'ambivalence largement décrits dans la littérature en psychologie et en marketing. Pour intégrer cette approche alternative de bidimensionnalité de l'attitude, il nous faut reconsidérer la structure des évaluations des consommateurs. Plus précisément, nous proposons de passer d'une conception bipolaire des évaluations (oscillant du négatif au positif) à une conception « quadripolaire ». Dans cette perspective, la positivité s'oppose à l'absence de positivité. Au contraire, la négativité s'oppose à l'absence de négativité. Un détour par un carré sémiotique (Greimas et Courtès, 1993) permet de construire du sens autour de ces oppositions.

Le carré sémiotique que nous proposons pour caractériser les réactions évaluatives des consommateurs est basé sur l'opposition entre « négatif » et « positif ». Deux concepts supplémentaires sont générés à partir de cet antagonisme : le « non négatif » et le « non positif ». La combinaison de ces termes permet de produire des méta-concepts issus d'une relation de contrariété : quand un consommateur évalue un produit ou service à la fois positivement et négativement, on parle d'*ambivalence*. Au contraire, un consommateur qui

ne perçoit un produit ou un service ni négativement, ni positivement est *indifférent*. La figure 6 représente le carré sémiotique que nous proposons pour décrire les évaluations des consommateurs.

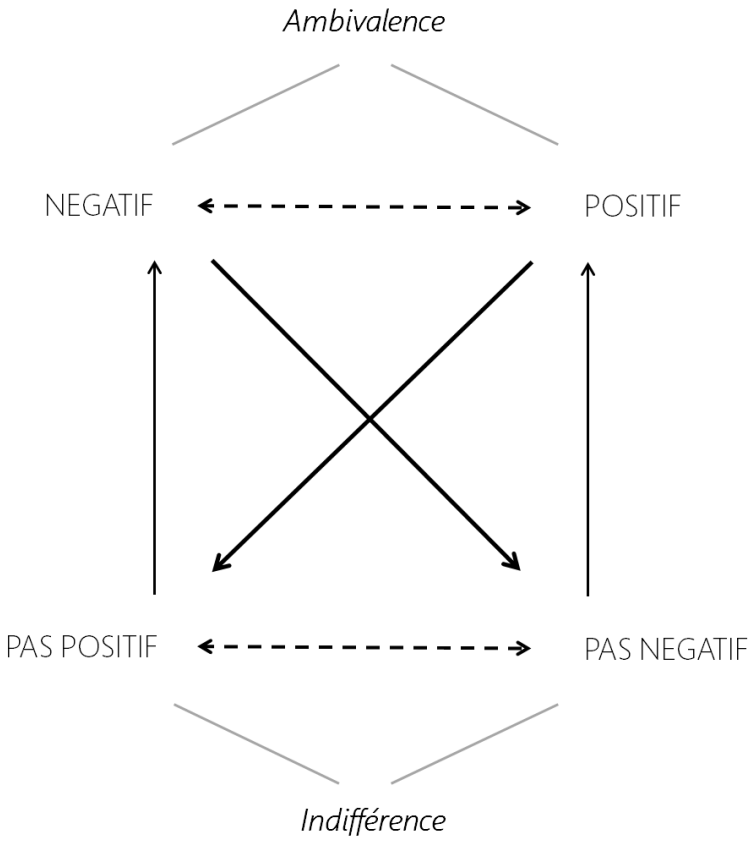


Figure 6. Le carré sémiotique des évaluations de consommateurs

Contrairement au classique continuum bipolaire qui envisage trois types d'évaluations (négatives, neutres et positives), le carré sémiotique nous permet d'affiner l'analyse en distinguant des évaluations ambivalentes et indifférentes.

Avant d'aller plus loin, nous avons réalisé une étude préliminaire visant à questionner les pratiques évaluatives mises en œuvre aujourd'hui sur les sites de commerce en ligne pour voir si les problèmes de mesures que nous avons soulevés sont toujours d'actualité.

Section 3. Etude préliminaire sur la mesure des évaluations globales de consommateurs dans le e-commerce

Cette étude préliminaire a pour double objectif de cerner les pratiques des sites de e-commerce en termes d'évaluations de consommateurs (1.) et de voir comment elles sont interprétées par les internautes (2.).

1. Panorama des outils utilisés dans le e-commerce pour mesurer et communiquer les évaluations globales de consommateurs

La mesure des évaluations globales de consommateurs est une réalité courante sur les sites de e-commerce. Néanmoins, force est de constater que les pratiques sont très diverses. Nous proposons ici de dresser un panorama des différents outils utilisés par les sites de vente en ligne pour mesurer et communiquer les évaluations globales de leurs clients. Nous allons tout d'abord décrire la méthode mise en œuvre pour collecter les données (1.1.), puis nous présenterons l'analyse et les résultats obtenus (1.2.), pour terminer par une brève conclusion (1.3.).

1.1. Méthode

Cette étude descriptive se focalise sur les principaux sites de commerce en ligne fréquentés par les internautes français. Pour constituer l'échantillon, nous avons consulté le « Top 15 du e-commerce » établi par la FEVAD (Fédération e-commerce et vente à distance) pour le troisième trimestre 2013. Ce classement a été effectué sur la base du nombre de visiteurs uniques par mois. A partir de ce palmarès, nous avons retiré certains sites de l'échantillon dans la mesure où ils ne proposaient pas de dispositif d'évaluation ou parce que ceux-ci prenaient une forme très éloignée des différentiels sémantiques. En effet, notre objectif étant de vérifier que la question du point central des différentiels sémantiques est toujours d'actualité, il était nécessaire de nous focaliser uniquement sur les outils de mesures comparables à celui-ci. Au final, 10 sites sur les 15 du palmarès ont été retenus. Le tableau 6 décrit précisément l'échantillon et les raisons d'exclusion des sites ayant été écartés.

Les 15 sites de e-commerce les plus visités (FEVAD)	Justification de l'intégration ou non dans l'échantillon (X : sites présents dans l'échantillon)
Amazon.fr	X
Cdiscount.com	X
Fnac.com	X
Ebay.fr	*Pas de dispositif d'évaluation des produits vendus. Possibilité d'évaluation des acheteurs et des vendeurs sous la forme d'un pourcentage d'évaluations positives et non d'une échelle.*
Voyages-sncf.fr	*Pas de dispositif d'évaluation des services vendus*
LaRedoute.fr	X
PriceMinister.com	X
Carrefour.fr	X
RueduCommerce.fr	X
Groupon.fr	*Pas de dispositif d'évaluation des services vendus*
Vente-privée.com	*Pas de dispositif d'évaluation des produits et services vendus*
Booking.com	*Le dispositif de mesure utilisé ne prend pas la forme d'une échelle. Il s'agit d'une note moyenne communiquée directement sous la forme d'un résultat.*
LeroyMerlin.fr	X
Darty.com	X
Auchan.fr	X

Tableau 6. Composition de l'échantillon de sites de e-commerce de l'étude préliminaire

Le recueil de données sur ces sites a été focalisé à la fois sur la forme prise par les dispositifs de mesure lorsqu'ils sont soumis aux clients souhaitant déposer une évaluation, ainsi que sur la manière dont les notes sont ensuite restituées aux internautes lecteurs de ces évaluations. Plus précisément, pour chaque site de l'échantillon, nous avons collecté les informations suivantes :

- Présentation synthétique du dispositif de mesure tel qu'il apparait aux internautes effectuant une recherche sur le site : forme et nombre de points composant l'échelle, présence ou absence d'étiquettes sémantiques associées à ces points ;
- Présentation détaillée de la note lorsque le client intéressé par l'un des résultats de sa recherche clique sur le lien pour obtenir de plus amples informations ;

- Dispositif de mesure proposé aux clients désireux de laisser une évaluation sur un produit ou un service consommé : forme et nombre de points composant l'échelle, présence ou absence d'étiquettes sémantiques associées à ces points.

1.2. Analyse et résultats

Nous avons cherché à déterminer si les outils de mesures utilisés sur les sites de e-commerce retenus dans notre échantillon sont bipolaires ou unipolaires. Pour ce faire, nous avons analysé le contenu des éventuelles étiquettes sémantiques qualifiant les points des échelles proposées. L'analyse descriptive effectuée est synthétisée dans le tableau 7.

La prédominance des échelles composées de 5 étoiles

Premier constat, les échelles composées de 5 étoiles prédominent largement. En effet, lorsqu'il s'agit de communiquer la note moyenne obtenue par un produit ou un service, c'est le format qui est systématiquement utilisé par les sites que nous avons étudiés. Le remplissage des étoiles à hauteur de la note moyenne obtenue, offre une visualisation concrète de celle-ci. Par ailleurs, ces échelles sont parfois accompagnées du nombre d'avis constituant la note et du détail chiffré du score obtenu sur cinq. Lorsqu'il s'agit pour un évaluateur de laisser une note, on retrouve à nouveau presque systématiquement les mêmes dispositifs composés de 5 étoiles. Notons toutefois une exception : sur le site RueduCommerce.fr, l'échelle proposée aux évaluateurs est constituée de 6 ronds allant de 0 à 5. Il s'agit du seul site de l'échantillon offrant la possibilité de laisser une note équivalente à 0 étoile.

Y compris sur les sites français, c'est donc la norme évaluative anglo-saxonne des 5 étoiles qui prédomine. Ainsi, les traditions d'évaluation des différents pays tendent à s'effacer au profit des pratiques des géants de l'Internet américain. Il existe toutefois, en dehors de notre échantillon quelques exceptions de sites qui utilisent leur propre système d'évaluation, bien souvent calqué sur celui du système scolaire de leur pays. Par exemple, le site français LaFourchette.com repose sur un système de note sur 10, tout comme Booking.com, site néerlandais. De même, certains sites de e-commerce allemands proposent d'évaluer avec des notes allant de 1 à 4, 1 étant la meilleure note.

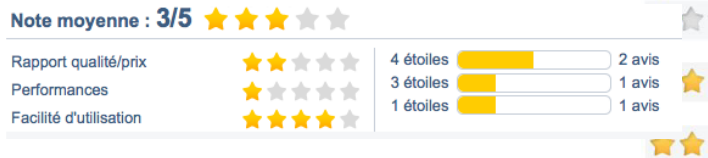
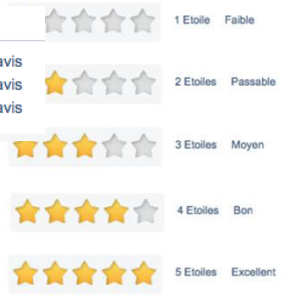
Après avoir cerné l'aspect général des dispositifs utilisés sur les sites de commerce en ligne, nous avons cherché à savoir dans quelle mesure ils pouvaient se rapprocher des différentiels sémantiques. Pour ce faire, nous avons examiné le contenu des étiquettes

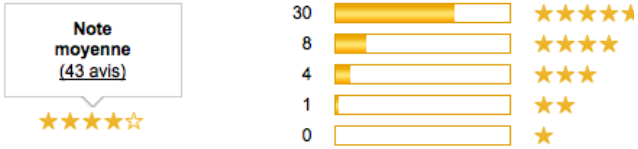
sémantiques associées aux différents points de ces échelles pour déterminer s'il s'agit de dispositifs unipolaires (allant de l'absence de réactions positives aux réactions très positives) ou bipolaires (allant du très négatif au très positif).

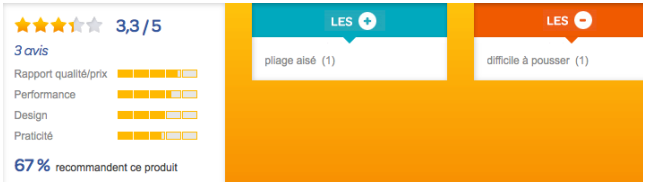
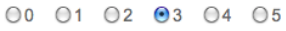
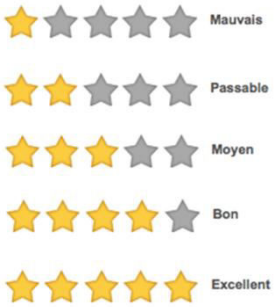
Des étiquettes sémantiques correspondant à des échelles bipolaires

Le premier constat concernant les étiquettes sémantiques est celui de leur absence pour les récepteurs de l'évaluation. En effet, lorsqu'un internaute est confronté par exemple à une note de 1/5 sur l'un des sites de notre échantillon, il ne dispose d'aucune indication pour savoir s'il s'agit d'une note négative ou d'une note faiblement positive. Bien sûr, dans le cas d'une note unique donnée par un seul évaluateur, il est possible de consulter le commentaire associé à celle-ci. Ce n'est pas le cas s'il s'agit d'une note globale résultant de la moyenne de plusieurs évaluations. Dans cette situation, l'interprétation des notes varie sans doute d'un internaute à l'autre.

La situation est différente lorsqu'un client souhaite laisser une évaluation dans la mesure où, dans la majorité des cas, les points de l'échelle sont alors accompagnés d'étiquettes sémantiques. Notons à cet égard que les termes employés diffèrent d'un site à l'autre. Trois sites de l'échantillon (Carrefour.fr, LeroyMerlin.fr et Auchan.fr) utilisent les qualificatifs suivants : 1 = « mauvais » ; 2 = « passable » ; 3 = « moyen » ; 4 = « bon » et 5 = « excellent ». Les termes employés par LaRedoute.fr et par Darty.com s'en rapprochent beaucoup. Pour ces cinq sites, l'étoile située au centre de l'échelle est associée au qualificatif « moyen », alors que les 2 étoiles de gauche sont associées à des termes négatifs et les 2 étoiles de droite à des termes positifs. Le site Amazon.fr oppose également le négatif au positif, mais c'est le terme « j'aime » qui caractérise la note de 3 étoiles sur 5. Il n'en demeure pas moins que ces sites utilisent des échelles bipolaires opposant de gauche à droite du continuum des notes négatives à des notes positives. Au contraire, les sites RueduCommerce.fr et Cdiscount.com proposent des dispositifs qui s'apparentent davantage à des échelles unipolaires. Ainsi, sur RueduCommerce.fr ce sont des chiffres qui sont associés aux points de l'échelle. Allant de 0 à 5, l'évaluation est unidirectionnelle (il n'y a pas de note négative). Sur Cdiscount.com les étiquettes sémantiques utilisées, allant de « faible » à « excellent » font du dispositif proposé une échelle unipolaire. Enfin, sur les sites Fnac.com et PriceMinister.com, il n'y a pas d'étiquette sémantique à destination des évaluateurs.

	Du point de vue du récepteur de l'évaluation		Du point de vue de l'évaluateur	Analyse du dispositif	
	Dispositif d'évaluation	Détail de la note ?	Etiquettes sémantiques ?	Outil bipolaire ou unipolaire ?	Point médian ?
Amazon	Echelle composée de 5 étoiles  Pas d'étiquette sémantique associée aux points		Oui, lorsque l'évaluateur passe sa souris sur la barre d'étoiles 	Bipolaire : oppose les réactions négatives aux réactions positives.	Oui, l'échelle étant composée de 5 points, la 3ème étoile se situe au centre de l'échelle. Toutefois, l'étiquette sémantique associe ce score à une note positive, qualifiée de « J'aime ».
Cdiscount	Echelle composée de 5 étoiles  Pas d'étiquette sémantique			Semble être plutôt unipolaire : on passe d'une faible réaction positive à une forte réaction positive	Oui, le point situé au centre de l'échelle est qualifié de « Moyen ».

	Du point de vue du récepteur de l'évaluation		Du point de vue de l'évaluateur	Analyse du dispositif	
	Dispositif d'évaluation	Détail de la note ?	Etiquettes sémantiques ?	Outil bipolaire ou unipolaire ?	Point médian ?
Fnac	Echelle composée de 5 étoiles ★★★★☆ Pas d'étiquette sémantique		Non, même échelle que celle visible par le récepteur de l'évaluation	On ne peut pas le savoir puisque les points ne sont pas caractérisés	Oui, l'échelle étant composée de 5 points, la 3ème étoile se situe au centre de l'échelle. Toutefois, elle n'est pas qualifiée.
La Redoute	Echelle composée de 5 étoiles ★★★★½ Pas d'étiquette sémantique			Bipolaire : oppose les réactions négatives aux réactions positives.	Oui, le point situé au centre de l'échelle est qualifié de « Moyen : pour dépanner ».
Price Minister	Echelle composée de 5 étoiles ★★★★☆ Pas d'étiquette sémantique	Pas de détail de la note	Non, même échelle que celle visible par le récepteur de l'évaluation	On ne peut pas le savoir puisque les points ne sont pas caractérisés	Oui, l'échelle étant composée de 5 points, la 3ème étoile se situe au centre de l'échelle. Toutefois, elle n'est pas qualifiée.

	Du point de vue du récepteur de l'évaluation		Du point de vue de l'évaluateur	Analyse du dispositif	
	Dispositif d'évaluation	Détail de la note ?	Etiquettes sémantiques ?	Outil bipolaire ou unipolaire ?	Point médian ?
Carrefour	Echelle composée de 5 étoiles  Pas d'étiquette sémantique			Bipolaire : oppose les réactions négatives aux réactions positives.	Oui, le point situé au centre de l'échelle est qualifié de « Moyen ».
Rue du Commer-ce	Echelle composée de 5 étoiles Pas d'étiquette sémantique	Pas de détail de la note		Semble être plutôt unipolaire puisqu'il n'y a pas de note négative	Pas de point central car l'échelle évaluateur est composée de 6 points.
Leroy Merlin	Echelle composée de 5 étoiles Pas d'étiquette sémantique			Bipolaire : oppose les réactions négatives aux réactions positives.	Oui, le point situé au centre de l'échelle est qualifié de « Moyen ».

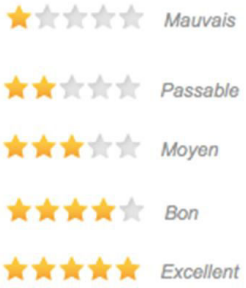
	Du point de vue du récepteur de l'évaluation			Du point de vue de l'évaluateur	Analyse du dispositif	
	Dispositif d'évaluation	Détail de la note ?		Etiquettes sémantiques ?	Outil bipolaire ou unipolaire ?	Point médian ?
Darty	Echelle composée de 5 étoiles 4/5 Pas d'étiquette sémantique				Bipolaire : oppose les réactions négatives aux réactions positives.	Oui, le point situé au centre de l'échelle est qualifié de « Moyen ».
Auchan	Echelle composée de 5 étoiles 4,5 sur 5 Pas d'étiquette sémantique				Bipolaire : oppose les réactions négatives aux réactions positives.	Oui, le point situé au centre de l'échelle est qualifié de « Moyen ».

Tableau 7. Analyse descriptive des dispositifs de mesure proposés sur les principaux sites de e-commerce français

Les dispositifs proposés aux évaluateurs sont donc majoritairement bipolaires, opposant les évaluations négatives aux évaluations positives. Toutefois, il ne s'agit pas de différentiels sémantiques au sens strict puisqu'il n'y a pas aux extrémités de l'échelle d'adjectif visible continuellement durant l'évaluation. En effet, les étiquettes sémantiques apparaissent seulement lorsque l'évaluateur passe sa souris sur les points du dispositif de mesure.

La question du point central

En adoptant massivement des échelles en 5 points, les dispositifs proposés par les sites de e-commerce soulèvent la question du sens attribué à l'étoile située au centre de l'échelle. Du point de vue des évaluateurs, ce point est généralement qualifié de « moyen ». Notons toutefois que le choix du point central correspond à une note de 3 sur 5, c'est-à-dire à une note située au-dessus de la moyenne. C'est également le cas lorsque, dans une étude marketing un répondant entoure par exemple le point 4 sur une échelle allant de 1 à 7. Néanmoins, la visualisation des notes par remplissage des points sur les sites de vente en ligne donne davantage l'impression de se situer au-dessus de la moyenne lorsque 3/5 étoiles sont colorées. L'étiquette sémantique « j'aime », utilisée par Amazon.fr pour qualifier le point central, va dans le sens de cette analyse. Quant aux récepteurs des évaluations, ils ne disposent d'aucune indication concernant l'interprétation du point central.

1.3. Conclusion

La prépondérance des échelles bipolaires en 5 points sur les sites de e-commerce renforce la nécessité de spécifier les réactions associées au point central. L'existence même de ce point est de nature à créer un flou interprétatif. En effet, celui-ci n'est jamais qualifié pour les récepteurs de l'évaluation et est le plus souvent caractérisé de « moyen » pour les évaluateurs. Or, une évaluation « moyenne » peut recouvrir différentes réalités. Pour appréhender plus précisément ce qui se cache derrière ce point central, nous avons complété cette étude par un focus sur l'interprétation faite par les consommateurs des différents points de ces échelles.

2. Sens attribué par les internautes aux points des dispositifs de mesure utilisés par les sites de e-commerce

Cette étude a pour objectif de clarifier l'interprétation faite par les lecteurs d'avis de consommateurs des différents points des échelles de mesure utilisées par les sites de e-

commerce. Nous souhaitons clarifier plusieurs points. Il s'agit tout d'abord de déterminer si les internautes perçoivent ces dispositifs comme unipolaires ou bipolaires. Ensuite, nous chercherons à savoir quels sont les qualificatifs associés aux différents points de ces échelles et plus précisément au point central. Cela sera également l'occasion de voir dans quelle mesure les étiquettes sémantiques proposées par les sites de e-commerce ont été intériorisées par les utilisateurs.

2.1. Méthode

Nous avons administré un questionnaire en ligne au laboratoire d'expérimentation de l'Université Paris-Dauphine dans le but de recenser les qualificatifs associés aux points d'un dispositif de mesure similaire à ceux utilisés par les sites de vente en ligne. L'étude a été réalisée auprès de 118 sujets ayant un âge moyen de 22,5 ans. Nous avons considéré que le fait de collecter les données auprès d'un échantillon jeune et estudiantin était un moyen indirect de toucher des individus familiers avec le e-commerce. L'étude a été focalisée sur une échelle composée de 5 étoiles puisqu'il s'agit du dispositif de mesure le plus répandu sur les sites de e-commerce. En effet, la fréquence importante de ce format de mesure nous a semblé propice à la restitution d'éventuelles connaissances accumulées sur le sens attribué aux différents points de ces échelles.

Après une brève introduction sur les pratiques d'évaluation dans le e-commerce, les sujets étaient invités à se projeter dans une situation d'évaluation d'un restaurant. Nous leur demandions ensuite d'indiquer l'adjectif ou l'expression qu'ils attribuaient à chaque point d'une échelle constituée de 5 étoiles (figure 7). La mise en contexte de l'évaluation d'un restaurant visait à limiter la variance des étiquettes sémantiques proposées.



Figure 7. Question posée pour déterminer le sens donné aux différents points des échelles

2.2. Analyse et résultats

Nous avons effectué deux types d'analyses à partir des données collectées : notre attention s'est tout d'abord portée sur la nature des dispositifs proposés par les répondants, puis sur les termes cités pour chaque point de l'échelle. Avant d'effectuer ces analyses, la base de données a été épurée. Sur les 118 lignes de réponses initiales, cinq ont été supprimées. En effet, un répondant ne semblait pas avoir compris la consigne et quatre autres ont proposé des qualificatifs se rapportant au niveau de gamme de restauration, plutôt qu'à l'appréciation d'un restaurant (par exemple « bas de gamme », « moyenne gamme » ou « luxueux »). Nous pensons que cette confusion est liée à l'interaction entre la mise en contexte de l'évaluation d'un restaurant et le format d'échelle choisi qui a pu être associé au système d'étoiles utilisé par le guide Michelin, par exemple. La base de données effective comptait donc 113 répondants.

Analyse de la nature des dispositifs proposés par les répondants

La première étape de notre analyse a consisté à déterminer la nature des échelles utilisées par notre échantillon. Dans le prolongement de cette analyse, nous avons tout d'abord distingué les échelles bipolaires des échelles unipolaires. A nouveau, nous préférons employer le terme « d'échelles bipolaires » à celui de « différentiels sémantiques » dans la mesure où les étiquettes sémantiques ne sont pas situées uniquement au niveau des extrémités de l'échelle comme cela devrait être le cas pour parler au sens strict de différentiel sémantique.

Pour rappel, une échelle est bipolaire lorsque les points situés à gauche du point central sont associés à des termes négatifs et que ceux situés à droite du point central sont associés à des termes positifs. Par exemple, nous avons classé la combinaison d'étiquettes sémantiques suivantes dans la catégorie des échelles bipolaires : 1 étoile = « très nul » ; 2 étoiles = « nul » ; 3 étoiles « moyen » ; 4 étoiles = « bon » et 5 étoiles = « très bon » (sujet n°5 de la base de données). 36% des réponses données ont été associées à la catégorie des échelles bipolaires.

Les échelles unipolaires se limitent une seule dimension du spectre évaluatif dont elles mesurent des variations en termes d'intensité. Dans le cas présent, les échelles que nous avons classées comme bipolaires comportaient des étiquettes sémantiques allant de la faible positivité à la forte positivité. Par exemple, le répondant n°48 de la base de données

a proposé la combinaison d'étiquettes sémantiques suivante : 1 étoile = « très faible » ; 2 étoiles = « faible » ; 3 étoiles « convenable » ; 4 étoiles = « bon » et 5 étoiles = « très bon ». 6% des combinaisons proposées ont été classées dans la catégorie des échelles unipolaires.

L'analyse a fait émerger une autre catégorie d'échelles que nous avons qualifiée « d'échelles bipolaires asymétriques ». Tout comme les échelles bipolaires, ces échelles opposent le négatif au positif, mais le point de passage entre les deux dimensions de l'évaluation n'est pas celui situé au centre de l'échelle. Sur ces échelles asymétriques, c'est la deuxième étoile qui marque la césure entre le négatif et le positif. Par exemple, le répondant n°41 de la base de données a proposé la combinaison d'étiquettes sémantiques suivante : 1 étoile = « mauvais » ; 2 étoiles = « moyen » ; 3 étoiles « bon » ; 4 étoiles = « très bon » et 5 étoiles = « excellent ». 58% des combinaisons proposées ont été classées dans la catégorie des échelles bipolaires asymétriques. La figure 8 résume cette analyse.

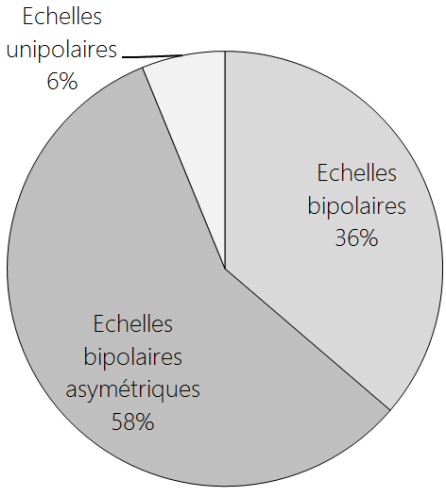
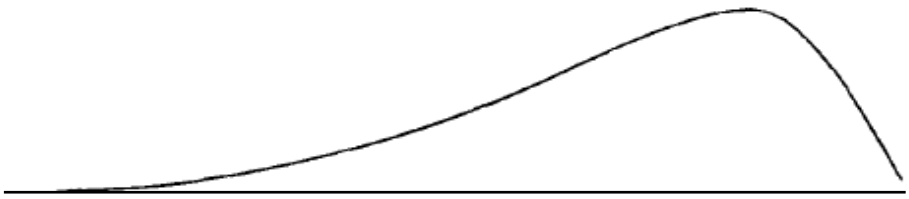


Figure 8. Analyse de la nature des dispositifs proposés par les répondants

La prédominance des échelles bipolaires asymétriques est intéressante. Cela souligne un décalage entre les étiquettes sémantiques employées par les sites de e-commerce lorsqu'un client souhaite laisser une évaluation et la perception qui en est faite par les récepteurs de l'évaluation. En effet, à l'exception du site Amazon.fr, aucun des sites étudiés dans la première partie de l'étude préliminaire n'utilisait ce type d'échelle. Lorsque l'information était disponible, le point central était qualifié de « moyen ». Ce décalage fait écho à la distribution asymétrique négative traditionnellement observée dans les enquêtes de satisfaction. En effet, les réponses évaluatives des consommateurs ne suivent pas une loi

normale et se caractérisent par une distribution asymétrique négative. Cela se traduit par des valeurs basses sur la partie gauche de la queue de la distribution, c'est-à-dire par un biais de positivité (Hurley et Estelami, 1998 ; Peterson et Wilson, 1992 ; Westbrook, 1980). Autrement dit, la plupart des personnes interrogées lors des enquêtes de satisfaction, se déclarent satisfaites ou très satisfaites (figure 9).



Extrême insatisfaction Extrême satisfaction

[*Source : Peterson et Wilson, 1992*]

Figure 9. Représentation graphique de la distribution asymétrique négative de la satisfaction

On retrouve très clairement cette distribution asymétrique négative dans les étiquettes sémantiques proposées par notre échantillon de répondants. Si dans le cadre des enquêtes de satisfaction, les individus se déclarent majoritairement satisfaits, dans cette étude également ils ont accordé une place plus importante à la satisfaction. En effet, dans 58% des cas les répondants ont associé la première étoile à un terme négatif, par exemple « mauvais », la deuxième étoile à un terme modéré, tel que « moyen », « bof » ou « pas terrible » et les trois étoiles suivantes ont été caractérisées par une gradation progressive de termes de plus en plus positifs allant de « bon » à « excellent » par exemple.

Comme nous l'avons souligné précédemment, cette asymétrie peut s'expliquer par le format du dispositif composé d'un nombre impair d'étoiles. En effet, le point que nous qualifions de « central » correspond à une note de trois sur cinq, c'est-à-dire à une note supérieure à la moyenne. Il n'est donc pas étonnant que les répondants lui aient majoritairement associé des qualificatifs « positifs ». Par ailleurs, la visualisation de la note choisie par remplissage favorise ce type de réponse.

Cette première étape de l'analyse, indiquant que les réponses se partagent à 94% entre des échelles bipolaires et des échelles bipolaires asymétriques, nous invite à explorer davantage le sens donné aux points centraux des échelles. En effet, contrairement à nos attentes, ce n'est pas systématiquement la troisième étoile qui constitue le point de partage

entre le positif et le négatif. Pour appréhender plus précisément l'interprétation des différents points de l'échelle, nous avons mené une analyse descriptive des étiquettes sémantiques proposées par notre échantillon.

Analyse des étiquettes sémantiques proposées pour chaque point de l'échelle

A partir des qualificatifs proposés pour chaque étoile, nous avons dressé un tableau sous Excel synthétisant le nombre d'occurrences associé à chaque descripteur. L'analyse a été effectuée séparément pour chaque point de l'échelle. Puis nous avons trié les données par ordre décroissant de manière à faire apparaitre les descripteurs les plus fréquents en haut du tableau. Nous avons ensuite effectué des regroupements suivant un critère de synonymie, de manière à synthétiser l'information dans un diagramme en secteur (figure 10). Par exemple, les termes « bien » et « bon » ont été regroupés. Nous avons particulièrement veillé à respecter l'intensité des qualificatifs employés lors de cette étape de regroupement. De manière à restituer un maximum d'information, nous avons essayé de limiter à 20% la catégorie « autres », constituée des termes n'ayant été proposés qu'une ou deux fois.

Le premier constat est que la variété des qualificatifs proposés est beaucoup plus faible pour les notes positives (4 et 5 étoiles), comme l'indiquent les digrammes en secteur de la figure 10. Ils se résument globalement aux termes « bon » ou « très bon » (22%) et « bien » ou « très bien » (59%) pour la quatrième étoile et à « excellent » (59%) pour la cinquième étoile. Lorsqu'il s'est agit de qualifier leur insatisfaction, les répondants ont été plus inspirés. Le terme « mauvais » a été proposé à hauteur de 42%. Il était suivi d'une variété de termes du registre de la recommandation (19%). Dans le cas présent, il s'agissait plus précisément de qualificatifs visant à déconseiller la fréquentation du restaurant : « à éviter », « déconseille », « non recommandé », « à ne pas fréquenter ». La présence du terme « médiocre », à hauteur de 8 % est intéressante, soulignant clairement le glissement sémantique du terme latin *medius* « qui est au milieu », vers le sens péjoratif qu'on lui attribue plus souvent actuellement.

Pour ce qui est de la deuxième et de la troisième étoile, la tâche de regroupement à été plus difficile à effectuer. En effet, le nombre de descripteurs uniques était bien plus important. En ce qui concerne la deuxième étoile, si les termes majoritairement cités étaient « moyen » et « bof » à hauteur de 21%, il y avait aussi des termes clairement négatifs, tels que « mauvais » (12%) ou « médiocre » (12%), ainsi que des étiquettes sémantiques

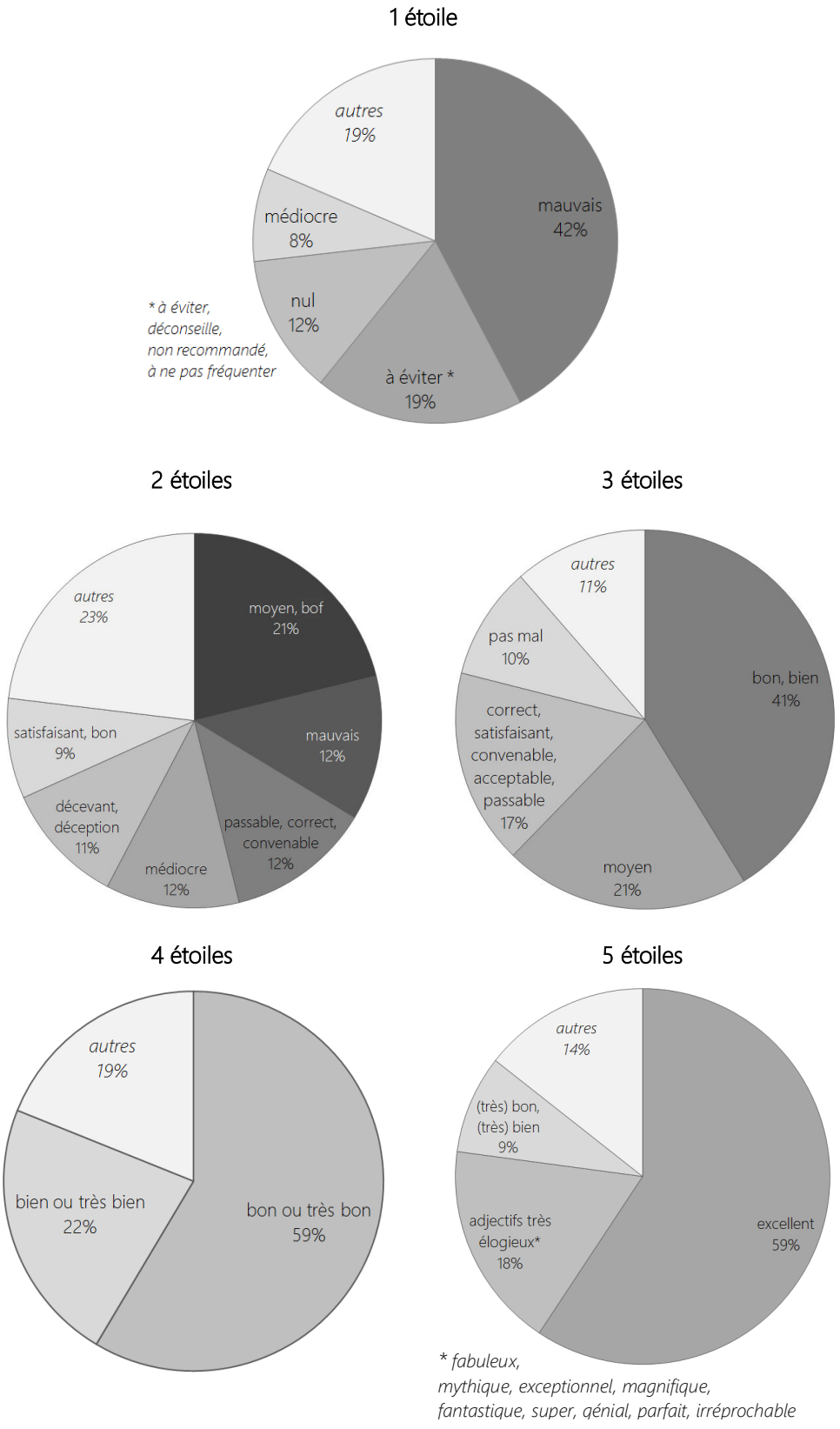


Figure 10. Analyse descriptive des étiquettes sémantiques proposées pour chaque point

positives comme « satisfaisant » et « bon » à hauteur de 9%. Il est également intéressant de citer les termes faisant référence à une comparaison avec les attentes : « passable », « correct », « convenable » (12%) ou encore « décevant » et « déception » (11%). Comme nous l'avions souligné lors de l'étape de description de la nature des dispositifs proposés, la troisième étoile a davantage été associée à des termes positifs. Les mots « bon » et « bien » étaient les plus fréquents, à hauteur de 41%. Arrivait ensuite le qualificatif « moyen », cité dans 21% des cas, suivi d'une variété de termes relevant de la comparaison aux attentes : « correct », « satisfaisant », « convenable », « acceptable » et « passable ». L'étiquette sémantique « pas mal », associée à une connotation légèrement positive, a également été citée dans 10% des cas.

Ces analyses sont représentées graphiquement dans la figure 10.

2.3. Conclusion

Cette analyse descriptive des étiquettes sémantiques proposées pour les différents points de l'échelle nous permet de réactualiser la problématique du point central des différentiels sémantiques notamment mise en évidence par Kaplan (1972). En effet, nous avions anticipé un phénomène d'uniformisation des « connaissances méthodologiques » des individus, liée à l'exposition des internautes à divers formats de dispositifs évaluatifs. Contrairement à nos attentes, une zone de flou sémantique se dessine au centre des échelles composées de 5 étoiles, pourtant très répandues sur les sites de e-commerce.

Cette étude préliminaire souligne avec encore plus d'acuité les limites des différentiels sémantiques pour saisir la part d'ambivalence contenue dans les évaluations de consommateurs. Plus précisément, nous établissons le double-constat suivant :

- La localisation précise des évaluations ambivalentes n'est pas claire : certains individus les situent au niveau de la deuxième étoile, d'autres au niveau de la troisième étoile des dispositifs de mesure les plus fréquemment utilisés sur les sites de e-commerce ;

- D'autre part, les évaluations ambivalentes sont noyées dans le groupe des réactions « moyennes ». Or, la polysémie du terme « moyen », n'est pas de nature à clarifier les choses. En effet, nous avons relevé au moins cinq significations

> concernant cet adjectif dans le *Petit Robert 2013*[4]. Cela va de la *neutralité* (« qui se trouve entre deux choses ») à la *modération* (« qui tient le milieu entre deux extrêmes »), en passant par le *fréquent* (« qui est le plus courant »), l'*indifférent* (« qui n'est ni bon, ni mauvais » ?) et enfin l'*ambivalent* (« que l'on établit en faisant une moyenne »… entre des réactions positives et des réactions négatives ?).

***

Ce dernier constat souligne à nouveau la nécessité de réfléchir à des solutions alternatives de mesure pour saisir la part d'ambivalence contenue dans les évaluations de consommateurs. Il apparaît en effet nécessaire de trouver un outil de mesure opérationnel qui permette de distinguer quatre pôles évaluatifs : les évaluations négatives, les évaluations indifférentes, les évaluations ambivalentes et les évaluations positives. C'est ce que nous proposons dans le chapitre 2 de cette thèse.

[4] *Ouvrage Collectif (2013), Le Petit Robert, Editions Le Robert, Paris.*

Chapitre 1	
Revue de la littérature	Le concept d'ambivalence La mesure de l'ambivalence
Etude préliminaire	La mesure des évaluations globales de consommateurs dans le e-commerce : pratiques et perception des internautes

↓

Chapitre 2	
Cadre conceptuel et hypothèses	Présentation de *l'Evaluative Space Grid* Résultats attendus

↓

Chapitre 3	
Etude 1	Validité convergente de *l'Evaluative Space Grid* avec les paires de continuums unipolaires
Etude 2	Validité convergente de *l'Evaluative Space Grid* avec les différentiels sémantiques, les discours des consommateurs et une mesure de l'ambivalence subjective

↓

Chapitre 4	
Etude 3	Test de *l'Evaluative Space Grid* comme outil de mesure de la satisfaction client

↓

Chapitre 5	
Etude 4	Etude qualitative après d'un échantillon d'experts : - qu'en pensent les chercheurs ? - qu'en pensent les professionnels ?

↓

Chapitre 6	
Etude 5	Place de l'incertitude dans les évaluations effectuées avec *l'Evaluative Space Grid*

– Chapitre 2 –

L'*Evaluative Space Grid*, un outil de mesure alternatif pour saisir la part d'ambivalence des consommateurs[5]

Nous avons présenté dans le chapitre précédent les limites des outils de mesure existants pour capturer la cooccurrence de réactions positives et négatives propre à l'ambivalence :

- Les différentiels sémantiques les fusionnent au niveau du point central avec les réactions indifférentes ;

- Les mesures proposant de faire appel à deux continuums unipolaires sont difficilement opérationnelles dans le cadre d'études marketing. En effet, elles supposent un déploiement méthodologique lourd car elles requièrent au minimum deux items pour la proposition de Kaplan (1972), voire quatre items si l'on suit les instructions de Russell et Carroll (1999). Compte tenu de l'enjeu des temps de passation des questionnaires, ce type de méthode reste limité aux recherches plus académiques qu'opérationnelles.

Nous présentons dans ce deuxième chapitre un outil de mesure bidimensionnel issu des recherches en psychologie (Section 1.) qui pourrait être adapté à la mesure des évaluations

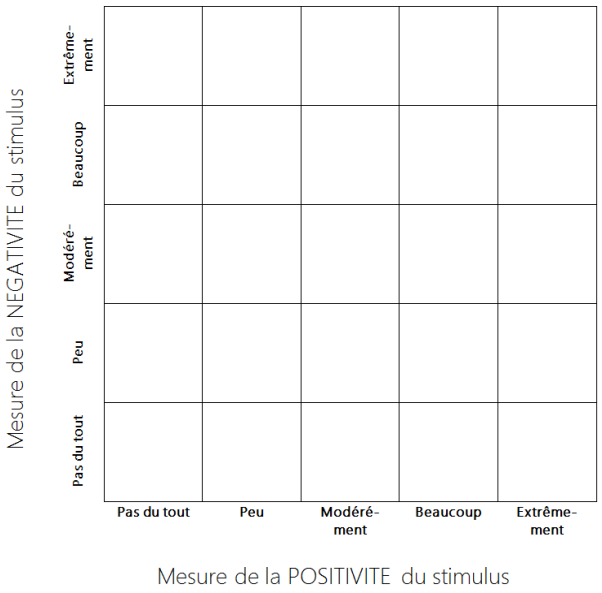
[5] *Cette étude a bénéficié du support de la Chaire AFNOR Performance des organisations de la Fondation Paris-Dauphine.*

de consommateurs dans le but de résoudre la question du point central des différentiels sémantiques (Section 2.).

Section 1. L'Evaluative Space Grid

1. Présentation de l'outil

L'*Evaluative Space Grid* (ou "grille de Larsen" en référence au premier auteur ; Larsen & al., 2009) est une mesure mono-item de la positivité et de la négativité d'un stimulus qui prend la forme d'une matrice 5 x 5 (figure 11). Elle a été développée par Larsen, Norris, McGraw, Hawkley et Cacioppo (2009) pour mesurer les réponses émotionnelles des individus lorsqu'ils sont confrontés à une large variété de stimuli visuels, auditifs et textuels. Sur cette grille, les répondants doivent cocher l'une des 25 cellules décrivant leur niveau de négativité (axe vertical allant de « pas du tout négatif » à « extrêmement négatif ») et leur niveau de positivité (axe horizontal allant de « pas du tout positif » à « extrêmement positif ») à l'égard d'un stimulus.



[*Source : Larsen et al. 2009*]

Figure 11. Traduction de l'*Evaluative Space Grid*

L'*Evaluative Space Grid* correspond simplement à la transformation du continuum linéaire que constitue le différentiel sémantique, en un plan bivarié. En effet, la forme matricielle de

la grille est obtenue en effectuant une translation verticale de la partie du différentiel sémantique consacrée à la mesure de la négativité (figure 12).

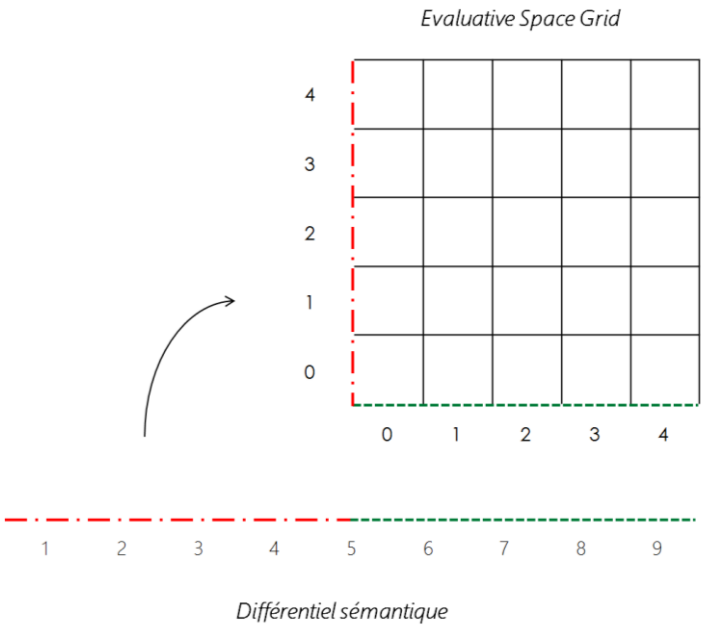


Figure 12. Du différentiel sémantique à l'*Evaluative Space Grid*

Larsen et ses collègues (2009) ont démontré dans le cadre d'une étude exploratoire sur ordinateur que les évaluations effectuées avec cette grille ne sont pas affectées par l'emplacement de l'origine du plan bivarié (le point de coordonnées 0 ; 0), qu'elle se situe en bas à gauche comme c'est le cas dans la figure 11 ou en haut à gauche. Ils ont également montré qu'il est possible d'intervertir l'axe de la positivité et l'axe de la négativité sans que cela n'impacte les réponses données. Cela confirme les résultats d'Andrade et Cohen (2007) qui avaient montré, en utilisant un format de mesure similaire dans le cadre d'une étude portant sur la co-activation des émotions positive et négative lors du visionnage de films d'horreur, que l'intitulé des axes de la positivité et de la négativité peut être interverti.

Toutefois, lorsque la grille est administrée sur papier, les résultats sont légèrement différents. Plus précisément, dans ce contexte, lorsque le point correspondant à l'origine de la grille se situe en bas à gauche de l'outil, les notes tendent à être plus élevées sur l'axe des ordonnées, c'est-à-dire sur l'axe de la négativité. En effet, dans le cadre d'une passation de questionnaire en ligne, l'emplacement du curseur peut être contraint au niveau de l'origine de la grille avant chaque évaluation, ce qui n'est pas réalisable dans le

cadre d'une passation sur papier. L'explication proposée par Larsen et ses collègues (2009) est que les répondants abordent alors l'outil suivant leurs habitudes de lecture, c'est-à-dire de gauche à droite et de bas en haut pour un français ou un américain. Ils auraient donc tendance à placer leur point d'ancrage au niveau de la cellule située en haut à gauche de la grille. Lorsque l'origine est située en bas à gauche, la cellule d'ancrage correspondrait alors aux réactions négatives maximales et positives minimales. C'est ce qui expliquerait pourquoi les réactions négatives sont légèrement plus élevées lorsque l'origine est située en bas à gauche plutôt qu'en haut à gauche de la grille.

2. Validité de l'outil

Larsen et ses collègues (2009) ont apporté la preuve de la validité convergente de l'*Evaluative Space Grid* avec deux autres mesures : une mesure effectuée en utilisant deux continuums unipolaires et une mesure effectuée suivant les préconisations de Russell et Carroll (1999), avec des d'échelles dichotomiques-puis-unipolaires. Pour rappel, les échelles dichotomiques-puis-unipolaires sont composées d'une paire de deux items : le premier est dichotomique (« ressentez-vous des réactions positives vis-à-vis de ce stimulus ? », réponse « oui » ou « non ») et si le sujet y répond par l'affirmative, le second item est unipolaire (« quelle est l'intensité des réactions positives éprouvées ? »). Le même dispositif est ensuite déployé pour mesurer les réactions négatives. Dans l'expérience de Larsen et ses collègues (2009), les participants ont pris part à des simulations de *coin flip*[6] dans lesquelles l'espérance de gain de chaque partie était nulle, les probabilités de gains et de pertes étant égales. Chaque essai pouvait se solder par un gain fort ou faible ou bien par une perte forte ou faible. Les gains forts faisaient office de stimuli positifs, les pertes fortes de stimuli négatifs, tandis que les gains et pertes faibles servaient de stimuli ambivalents. En effet, pour ces stimuli, le sujet s'attendait à gagner (dans la condition gains faibles) ou perdre (dans la condition pertes faibles) une somme supérieure à ce qu'il recevait effectivement, ce qui avait pour objectif de susciter un mélange de réactions positives et négatives.

Les résultats indiquent que l'*Evaluative Space Grid* offre une mesure valide de la positivité et de la négativité. Plus précisément, tout comme les mesures effectuées avec deux continuums unipolaires, la grille de Larsen permet de capturer les réactions ambivalentes

[6] *Le « coin flip » est un terme générique anglais désignant tous les jeux d'argent dans lesquels le joueur a une espérance de gain nulle. Il peut s'agir par exemple du jeu de « pile ou face » ou « chiquenaude sur une pièce ».*

associées aux gains et aux pertes moins importants que prévus. De plus, la corrélation entre les notes obtenues avec chaque outil était au moins égale à 0,97. Enfin, l'utilisation de l'*Evaluative Space Grid* a permis de réduire le temps de passation du questionnaire de 27% par rapport aux paires de continuums unipolaires et de 47% par rapport à la technique des échelles dichotomiques-puis-unipolaires. Dans le cadre d'une étude comportant 200 stimuli, les auteurs estiment que leur outil permettrait de gagner respectivement 5 et 10 minutes par rapport aux paires de continuums unipolaires ou aux échelles dichotomiques-puis-unipolaires.

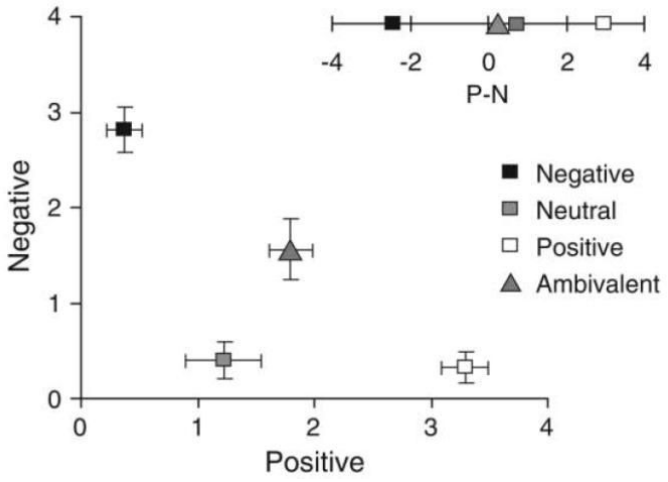
Ils ont également démontré que l'*Evaluative Space Grid*, à la différence des différentiels sémantiques, permet de distinguer les attitudes ambivalentes des attitudes indifférentes. Pour ce faire, ils ont demandé à 18 participants d'envoyer par mail divers inventaires. L'une des listes demandées consistait à énumérer dix choses à l'égard desquelles le répondant est ambivalent. Bill Clinton et la peine de mort étaient donnés en exemple. Quelques jours plus tard, les participants étaient invités à évaluer, en utilisant l'*Evaluative Space Grid*, une série de stimuli incluant les objets d'attitude ambivalents qu'ils avaient listés. Ces stimuli étaient accompagnés de stimuli négatifs (comme le terrorisme ou les embouteillages), positifs (comme les grandes vacances ou le beau temps) et indifférents (comme la laitue ou le papier-peint). Les résultats montrent que la grille de Larsen permet de capturer la distinction caractérisant la différence entre les attitudes ambivalentes et indifférentes. En effet, comme l'indique la figure 13 :

- les stimuli positifs ont donné lieu à des évaluations fortement positives (avec une note moyenne d'environ 3,2/4 sur l'axe de la positivité) et faiblement négatives (avec une note moyenne d'environ 0,2/4 sur l'axe de la négativité) ;

- les stimuli négatifs ont donné lieu à des évaluations faiblement positives (avec une note moyenne d'environ 0,2/4 sur l'axe de la positivité) et fortement négatives (avec une note moyenne d'environ 2,8/4 sur l'axe de la négativité) ;

- les stimuli neutres ou indifférents ont donné lieu à des évaluations faiblement positives (avec une note moyenne d'environ 1,2/4 sur l'axe de la positivité) et faiblement négatives (avec une note moyenne d'environ 0,3/4 sur l'axe de la négativité) ;

- les stimuli ambivalents ont donné lieu à des évaluations modérément positives (avec une note moyenne d'environ 1,9/4 sur l'axe de la positivité) et modérément négatives (avec une note moyenne d'environ 1,5/4 sur l'axe de la négativité).



[*Source : Larsen et al. 2009*]

Figure 13. Localisation des évaluations des différents types de stimuli sur la grille de Larsen

Les auteurs ont également mis en évidence la validité convergente de l'*Evaluative Space Grid* avec les différentiels sémantiques. Pour ce faire ils ont calculé la différence nette entre la note de positivité et la note de négativité à partir des axes de leur outil et ils ont montré qu'elle était très corrélée avec les scores obtenus sur des différentiels sémantiques (voir continuum situé en haut à droite sur la figure 13).

Cette étude présente néanmoins quelques limites. D'une part, elle a été réalisée sur un échantillon de petite taille (18 participants). D'autre part, dans la mesure où les stimuli ambivalents avaient été désignés comme tels, quelques jours avant la tâche d'évaluation par les participants, il est impossible de savoir de quel type d'objet d'attitude il s'agissait. Notre objectif est donc d'explorer le potentiel de cet outil pour saisir des évaluations de consommateurs dans le cadre du débat méthodologique concernant la distinction entre indifférence et ambivalence.

Section 2. Application de l'*Evaluative Space Grid* à la mesure des évaluations de consommateurs

Nous proposons dans cette thèse d'adapter l'*Evaluative Space Grid* à la mesure des évaluations globales de consommateurs. En effet, offrant une mesure mono-item des réactions positives et négatives en phase pré-achat ou post-achat, l'outil de Larsen et ses collègues devrait permettre de pallier les limites des outils existants. La figure 14 illustre comment la grille pourrait être utilisée par exemple pour mesurer la satisfaction client.

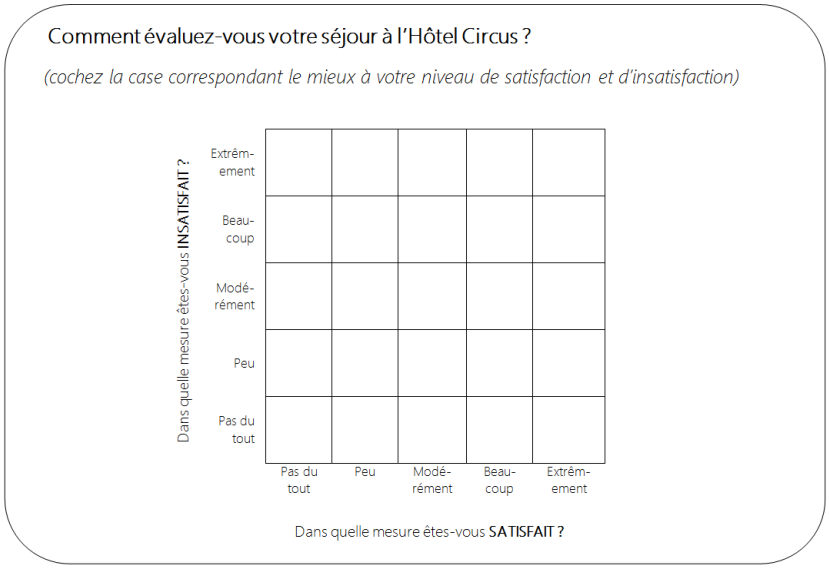


Figure 14. Application de l'*Evaluative Space Grid* à la mesure de la satisfaction

Le passage d'une conception unidimensionnelle à une conception bidimensionnelle nous permet de mettre en évidence quatre types de réactions évaluatives de consommateurs. Nous proposons la répartition suivante sur la grille (figure 15) :

- *Les réactions négatives* devraient se situer en haut à gauche de la grille, zone de l'outil associée à des réactions négatives « modérées » à « extrêmes » et à « pas du tout » ou « peu » de réactions positives ;

- A l'inverse, les *réactions positives* devraient se situer en bas à droite de la grille, zone de l'outil associée à des réactions positives « modérées » à « extrêmes » et à « pas du tout » ou « peu » de réactions négatives ;

- Les *réactions ambivalentes devraient* se situer en haut à droite de la grille, zone de l'outil correspondant à toutes les combinaisons de positivité et de négativité « modérée » à « extrême ».

- Enfin, les *réactions indifférentes* devraient se situer en bas à gauche de la grille, zone de l'outil correspondant aux combinaisons de « pas du tout » ou de « peu » de réactions positives et négatives.

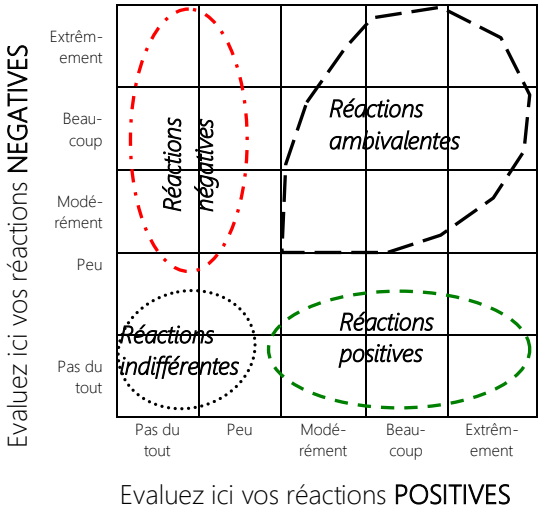


Figure 15. Localisation des quatre types de réactions évaluatives sur l'*Evaluative Space Grid*

1. Délimitation de l'usage de l'outil

Cette section vise à délimiter l'utilisation que nous proposons de faire de l'outil dans le cadre d'études consommateurs. En effet, nous envisageons l'*Evaluative Space Grid* comme une mesure des réactions affectives et cognitives (1.1.) mono-item (1.2.) globale (1.3.).

1.1. Une mesure des réactions affectives et cognitives

Le premier point de discussion concerne la délimitation des réactions mesurées par l'*Evaluative Space Grid*. Lorsque l'on se penche sur la manière dont Larsen et ses collègues (2009) envisagent leur outil, l'approche est assez généraliste. Il s'agit d'une « mesure mono-item de la positivité et de la négativité » permettant de saisir des « réactions évaluatives ». Toutefois, les termes « d'affect » ou « d'émotion » sont également présents sous leur plume, indiquant que l'*Evaluative Space Grid* s'inscrit au départ dans le camp des mesures affectives.

Ce premier constat ne doit pas nécessairement cantonner la grille à la mesure de réactions affectives. L'étude des *attitudes* ambivalentes recouvre un large spectre de réactions, incluant des réponses d'ordre cognitif. Nous positionnons donc l'*Evaluative Space Grid* comme un outil de mesure des réactions affectives et cognitives, répondant ainsi à l'injonction de Zaltman (1997) concernant la nécessité de développer des méthodes de recherche permettant de prendre en compte l'émotionnel et le rationnel.

1.2. Une mesure mono-item

Nous avons présenté l'*Evaluative Space Grid* comme une mesure *mono-item* des réactions positives et négatives permettant de saisir la part d'ambivalence des consommateurs. On pourrait toutefois argumenter qu'il suffirait d'utiliser une mesure multi-items pour saisir d'éventuelles réactions ambivalentes. Par exemple, dans la perspective d'une application de la grille de Larsen à la mesure de la satisfaction, une échelle de type SERVQUAL (Parasuraman, Zeithaml et Berry, 1988) permettrait de saisir des réactions positives sur certains attributs et négatives sur d'autres.

Nous ne pensons pas que l'*Evaluative Space Grid* soit à envisager comme un outil concurrent à de telles mesures. En effet, elles sont essentielles dès lors que l'on cherche à analyser de manière détaillée les performances d'un produit ou d'un service. La question de la mesure globale relève d'un registre différent. Plus précisément, elle s'inscrit dans la perspective de la recherche de parcimonie au niveau du nombre d'items. Il s'agit d'une préoccupation importante dans l'industrie des études et dans le cadre de la collecte d'avis de consommateurs en ligne, secteurs d'activité qui cherchent à lutter contre les abandons lors des passations de questionnaires. En ce qui concerne la mesure des avis de consommateurs en ligne, les professionnels se réfèrent par exemple à la règle des « deux minutes, cinq questions » (cf. Chapitre 5, Section 2). De telles pratiques soulignent la nécessité de proposer des mesures globales permettant de capturer un maximum d'informations en un minimum de temps.

1.3. Une mesure globale

Ensuite, nous proposons d'utiliser l'*Evaluative Space Grid* comme une mesure des réactions *globales* de consommateurs. En effet, la note obtenue sur cet outil résulte de la combinaison entre les notes données sur chaque axe. L'agrégat entre les deux notes globales est déterminé « mécaniquement » par la grille, avec un poids égal donné à

chaque note. La grille ne nous donne pas d'élément sur la composition du score de positivité et du score de négativité.

En ce sens, l'*Evaluative Space Grid* offre une approche différente de celle proposée par exemple par Llosa (1996) avec la mesure de l'asymétrie des contributions des attributs à la satisfaction globale. Si nous n'excluons pas que la grille de Larsen puisse être utilisée comme support à une telle approche dans de futures recherches, ce n'est pas l'objectif de cette thèse. Nous souhaitons positionner l'*Evaluative Space Grid* comme un outil offrant une sophistication par rapport au différentiel sémantique, dans la mesure où il permet de distinguer les réactions globales positives et globales négatives, sans entrer dans le processus de contribution des attributs à la satisfaction globale.

2. Résultats attendus

Comme nous l'avons vu précédemment, dans le cadre d'une évaluation globale supposant l'intégration des différents aspects de l'objet à évaluer, les différentiels sémantiques fusionnent les réactions positives et négatives au niveau du point central du continuum. En d'autres termes, cet outil incite les répondants à condenser l'information évaluative disponible dans leur esprit en les forçant à la restituer le long d'un continuum bipolaire. Cela entraine une perte d'information.

Nous pensons que l'*Evaluative Space Grid* va améliorer la validité de construit des évaluations de consommateurs dans la mesure où elle va permettre de déployer sur deux dimensions toute l'information disponible dans l'esprit des consommateurs. Plus précisément, cela devrait passer par les mécanismes suivants :

- Avant tout, par la **distinction entre indifférence et ambivalence** permettant de restituer l'intégralité du spectre évaluatif.

- Ensuite, par la **capture de plusieurs degrés d'ambivalence**. Comme nous l'avons expliqué précédemment, même lorsque le point central des différentiels sémantiques est associé à un terme autorisant les réponses ambivalentes (par exemple : « à la fois positif et négatif »), le problème n'est pas totalement résolu dans la mesure où cela revient à envisager l'ambivalence comme une variable nominale binaire. Dans cette perspective, si le répondant choisit le point central il est ambivalent, sinon il ne l'est pas. En d'autres termes, cela exclut toute variation

d'intensité de l'ambivalence. Pourtant, Heuvinck (2012) établit une différence entre l'ambivalence manifeste et l'ambivalence anticipée. Les réactions d'ambivalence manifeste correspondent à des situations dans lesquelles les personnes sont capables d'identifier clairement les points positifs et négatifs qui composent leur évaluation. En revanche, l'ambivalence anticipée fait référence au fait d'avoir à l'esprit de manière consciente un certain nombre d'informations positives (négatives) à l'égard de l'objet évalué et d'anticiper d'éventuelles informations négatives (positives). Il y aurait donc différents « degrés d'ambivalence » qui pourraient se répartir le long de la diagonale de la grille.

Cette amélioration de la validité de construit devrait également se traduire par une amélioration de la validité prédictive de la mesure.

En effet, dans le domaine des sciences politiques, des recherches ont montré que les conséquences comportementales de l'indifférence et de l'ambivalence ne sont pas les mêmes. Plus précisément, il a été mis en évidence que les citoyens ambivalents sont plus enclins à se rendre aux urnes que les citoyens indifférents (Yoo, 2010). De même, les personnes indifférentes à la politique sont moins promptes à s'impliquer dans des activités politiques ou des actions de soutien à un parti que les personnes ambivalentes (Thornton, 2011).

Ces recherches sont cohérentes avec les travaux en marketing et psychologie indiquant que l'ambivalence peut se traduire par des comportements d'approche, lorsque les individus cherchent à réduire le conflit qu'elle suscite en eux (Jonas, Diehl et Brömer, 1997 ; Maheswaran et Chaiken, 1991 ; Sengupta et Johar, 2002 ; Srull et Wyer, 1989).

Dans la mesure où l'indifférence et l'ambivalence sont associées à des comportements différents dans la littérature, nous pensons que la distinction de ces deux types d'évaluations grâce à l'*Evaluative Space Grid* devrait permettre d'améliorer la prédiction des comportements à partir des évaluations de consommateurs.

***

Après avoir présenté l'*Evaluative Space Grid*, délimité son usage et développé les résultats attendus de son utilisation, la suite de notre travail consistera à mettre à l'épreuve ce nouvel outil pour saisir la part d'ambivalence contenue dans les évaluations de consommateurs

– Introduction de la partie empirique –

Dans le chapitre 2 nous avons présenté L'*Evaluative Space Grid*, un nouvel outil de mesure susceptible d'améliorer la mesure des évaluations globales de consommateurs en prenant en compte la part d'ambivalence qu'elles contiennent.

Avant de déployer ce nouvel outil dans la conduite d'études consommateurs, il est nécessaire de préciser les modalités retenues pour sa mise à l'épreuve dans le cadre de la thèse. En effet, cinq études étant mobilisées pour ce test, nous avons décidé de nous limiter à un domaine d'application et à une opérationnalisation de la grille de Larsen afin de limiter les sources de variances externes et de faciliter la mise en perspective des résultats entre les différentes études.

Une application à l'évaluation des services

Nous avons décidé d'explorer l'intérêt de l'*Evaluative Space Grid* en l'appliquant à l'évaluation des services pour plusieurs raisons.

Tout d'abord, Oliver (1997) souligne que les services offrent des propriétés propices à l'étude de l'ambivalence. En effet, parce qu'il s'agit d'objets d'évaluation complexes, souvent composés d'une multitude d'attributs, ils présentent une plus grande probabilité de cumuler à la fois des attributs évalués positivement et négativement. Notre objectif étant

d'isoler la part d'ambivalence contenue dans les évaluations de consommateurs, il semble plus pertinent d'utiliser la grille de Larsen pour évaluer une expérience au restaurant, un service bancaire ou une prestation de coiffure que pour apprécier une bouteille de Coca Cola ou une table basse.

Par ailleurs, il ressort de la revue de la littérature que l'ambivalence à l'égard des services a été peu étudiée. A l'exception de la recherche de Clarkson et ses collègues {, 2008 #144} portant les supermarchés, ce domaine d'application offre de multiples opportunités de contribution.

Une opérationnalisation similaire à celle de Larsen et ses collègues

Dans la mesure où notre travail constitue la première exploration de l'*Evaluative Space Grid* dans la conduite d'études consommateurs, nous avons décidé de rester le plus fidèle possible au travail de Larsen et ses collègues (2009). En effet, avant de réfléchir à d'éventuelles adaptations de ce nouvel outil, il est important de savoir comment les répondants se l'approprient dans sa version originelle. Nous avons donc traduit l'outil de Larsen et ses collègues et l'avons programmé sous Wysuforms. Cette transposition a été accompagnée des adaptations suivantes :

- Modalités de réponses sous forme de points. Nous avons abandonné les cellules de l'*Evaluative Space Grid* au profit de points du fait de contraintes techniques propres au logiciel utilisé pour collecter les données. Cette adaptation présente toutefois l'avantage de proposer un format de réponses comparable à celui des différentiels sémantiques.

- Noms des axes et des étiquettes sémantiques associées aux points des axes en fonction des variables mesurées. Pour ce qui est des étiquettes sémantiques des axes, nous avons privilégié des expressions courtes comme « opinions négatives », aux questions utilisées par Larsen et ses collègues (voir figure 16) pour faciliter la lecture de l'outil. A nouveau, cette adaptation présente l'avantage d'une comparaison plus immédiate avec les étiquettes situées aux extrémités des différentiels sémantiques. En ce qui concerne les termes associés aux points des axes, nous avons opté pour des étiquettes sémantiques très proches de celles proposées par Larsen et ses collègues (2009), même si quelques adaptations à la marge ont été opérées pour correspondre aux variables mesurées (voir figure 16).

- **Origine de la grille située en bas à gauche.** L'emplacement du point de coordonnées [0 ; 0] pose question. Larsen et ses collègues (2009) ont mis en évidence en effet qu'il affecte le sens de lecture des répondants et, ce faisant, leurs évaluations. Plus précisément, ils ont montré que les évaluations rapportées sur l'axe de la négativité sont légèrement plus élevées lorsque l'origine est située en bas à gauche de la grille et qu'il n'est pas possible de contraindre l'emplacement du curseur des répondants au niveau de l'origine de l'outil avant chaque évaluation (par exemple dans le cas d'une collecte papier) (cf. Chapitre 2, Section 1). En même temps, de nombreuses recherches ont mis en évidence l'existence d'un biais de positivité dans la mesure de la satisfaction (Hurley et Estelami, 1998 ; Peterson et Wilson, 1992), ce que notre étude confirme. Nous pensons qu'inciter les répondants à assumer la part de négativité contenue dans leur évaluation peut offrir un moyen de corriger ce biais de positivité. Pour cette raison, nous avons choisi, comme Larsen et ses collègues (2009) de positionner l'origine de l'outil en bas à gauche de la grille, sans toutefois contraindre le curseur à cet emplacement.

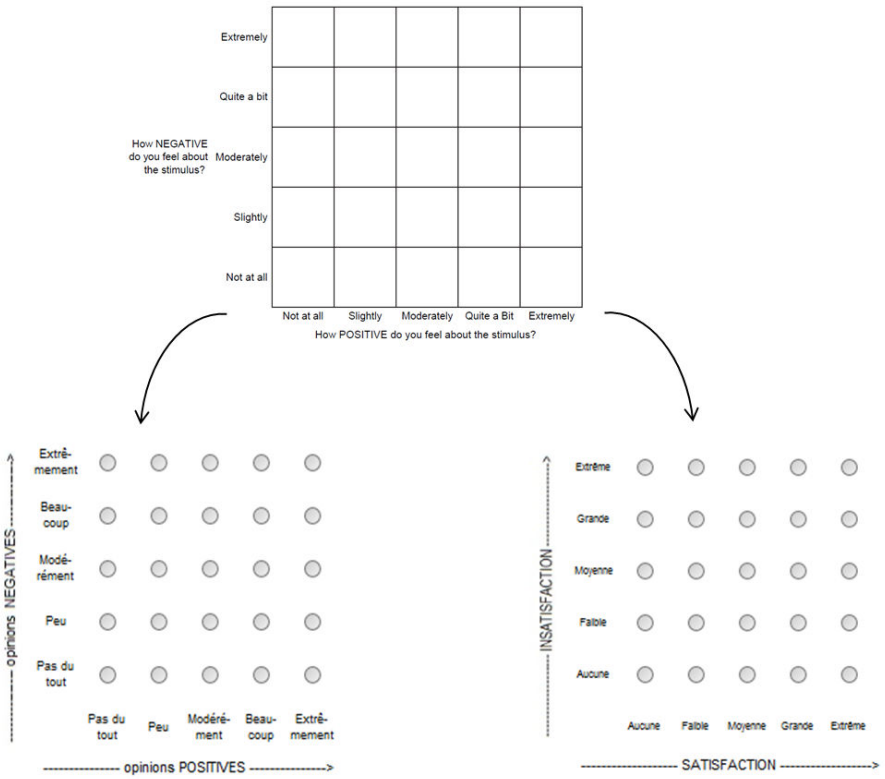


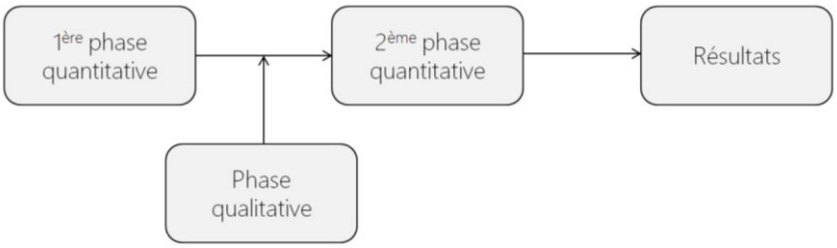
Figure 16. Opérationnalisation de l'*Evaluative Space Grid* pour la mesure des opinions (à gauche) et de la satisfaction globale (à droite)

- **Réactions positives en abscisses et réactions négatives en ordonnées.** Larsens et ses collègues (2009) ont démontré que l'inversion des étiquettes sémantiques des axes n'affecte pas les résultats. Nous avons décidé de placer la mesure des réactions négatives en ordonnées parce que cela nous semble être l'opérationnalisation la plus proche du différentiel sémantique. En effet, comme l'indique la figure 12 (page 61), dans cette configuration, le passage d'un différentiel sémantique à une *Evaluative Space Grid* correspond alors à une simple translation verticale de la partie consacrée à la mesure de la négativité.

La figure16 représente la manière dont nous avons opérationnalisé l'*Evaluative Space Grid* en fonction des variables mesurées, à partir de la proposition originelle de Larsen et ses collègues (2009).

Un design mixte séquentiel

La partie empirique de la thèse est composée de cinq études empiriques organisées suivant adopte un design mixte séquentiel (Creswell, 1995, p. 177). Plus précisément, elle est comporte une première phase quantitative qui a été suivie d'une phase qualitative visant à expliquer les résultats quantitatifs obtenus. Cette phase qualitative a fait apparaître de nouvelles pistes de recherche, testées ensuite dans une deuxième phase quantitative (figure 17).



[*Source : Ulin et al. 1996*]

Figure 17. Design mixte séquentiel de la recherche

Plus précisément, le plan de la partie empirique est le suivant :

Notre premier objectif était de savoir si l'*Evaluative Space Grid*, permettait de saisir la part d'ambivalence contenue dans les évaluations des consommateurs. Pour l'atteindre, les études 1 et 2 (Chapitre 3) sont consacrées au test de la validité convergente de la grille avec

les outils de mesure existants de l'ambivalence et des réactions évaluatives globales de consommateurs.

Notre deuxième objectif était de tester le pouvoir prédictif de l'*Evaluative Space Grid*. Pour l'atteindre, l'étude 3 (Chapitre 4) vise à démontrer empiriquement la supériorité psychométrique de la grille de Larsen dans le cadre d'une mesure de la satisfaction client.

Enfin, notre dernier objectif était de replacer l'*Evaluative Space Grid* au cœur des problématiques de mesure des études consommateurs. Une étude qualitative (Etude 4, Chapitre 5) est réalisée pour comprendre le sens donné par un panel d'experts aux différentes modalités de réponses de la grille. Celle-ci met notamment en évidence une zone d'incertitude au niveau du point central de la grille. Une étude quantitative la prolonge afin de tester empiriquement certaines des propositions qui en sont issues (Etude 5, Chapitre 6).

Le tableau 8 présente de manière synthétique le contenu et la méthode déployée dans les cinq études de la thèse.

Etudes	Objectif	Méthode	Echantillon	Analyses	Outils	Variables dépendantes
Etude 1	Test de la validité convergente de l'outil	Etude expérimentale	71 personnes	Test de différences de moyennes	-Grille de Larsen -Différentiel sémantique	Aucune
Etude 2	Test de la validité convergente de l'outil	Etude expérimentale	130 personnes	Test de différences de moyennes	-Grille de Larsen -Deux continuums unipolaires -Verbatims recodés -Mesure subjective de l'ambivalence	Aucune
Etude 3	Test de la validité de construit et de la validité prédictive de l'outil	Etude expérimentale	400 personnes	Equations structurelles	-Grille de Larsen -Différentiel sémantique	-Intentions de fidélité -Intentions de bouche-à-oreille
Etude 4	Compréhension de la signification associée aux points de l'outil	Entretiens semi-directifs	13 personnes	Analyse de contenu catégorielle thématique	Grille de Larsen	Aucune
Etude 5	Validation des propositions issues de la phase qualitative	Etude de satisfaction	200 personnes	Régressions	Grille de Larsen	-Implication -Confiance dans la réponse donnée

Tableau 8. Présentation des cinq études empiriques de la thèse

Chapitre 1	
Revue de la littérature	Le concept d'ambivalence La mesure de l'ambivalence
Etude préliminaire	La mesure des évaluations globales de consommateurs dans le e-commerce : pratiques et perception des internautes

↓

Chapitre 2	
Cadre conceptuel et hypothèses	Présentation de *l'Evaluative Space Grid* Résultats attendus

↓

Chapitre 3	
Etude 1	Validité convergente de *l'Evaluative Space Grid* avec les paires de continuums unipolaires
Etude 2	Validité convergente de *l'Evaluative Space Grid* avec les différentiels sémantiques, les discours des consommateurs et une mesure de l'ambivalence subjective

↓

Chapitre 4	
Etude 3	Test de *l'Evaluative Space Grid* comme outil de mesure de la satisfaction client

↓

Chapitre 5	
Etude 4	Etude qualitative après d'un échantillon d'experts : - qu'en pensent les chercheurs ? - qu'en pensent les professionnels ?

↓

Chapitre 6	
Etude 5	Place de l'incertitude dans les évaluations effectuées avec *l'Evaluative Space Grid*

– Chapitre 3 –

Validité convergente de l'*Evaluative Space Grid*

Ce chapitre vise à explorer la validité convergente de l'*Evaluative Space Grid*. La première étape de notre travail consiste à démontrer que les évaluations issues de la grille de Larsen convergent, c'est-à-dire qu'elles sont suffisamment corrélées, avec les évaluations issues des outils que nous entendons concurrencer.

Dans le cadre de l'étude 1 nous allons comparer les évaluations issues de l'*Evaluative Space Grid* à celles provenant de paires de deux continuums unipolaires (Section 1). L'étude 2 est focalisée sur les outils que la grille pourrait plus précisément concurrencer dans le cadre d'études consommateurs. Dans cette perspective, nous allons confronter les évaluations effectuées avec la grille de Larsen à celles provenant de trois dispositifs alternatifs : des discours réels de consommateurs, des évaluations obtenues sur des différentiels sémantiques et un score moyen d'ambivalence subjective (Section 2).

Section 1. Analyse de la validité convergente de la grille avec les paires de continuums unipolaires (Etude 1)

Cette étude est concentrée sur la mesure des réactions positives et négatives suscitées par la lecture de corpus d'avis de consommateurs en ligne dans le contexte de l'hôtellerie-

restauration. Nous avons opté pour cette application parce qu'elle offre la possibilité d'effectuer de libres combinaisons d'avis positifs et négatifs de manière à créer des corpus de commentaires de valence variable (négatif, positif et ambivalent en l'occurrence). Par ailleurs, dans une perspective de réalisme, la lecture d'avis de consommateurs se prête particulièrement bien à un recueil de données en ligne. L'objectif de cette étude est d'étendre à un contexte de consommation les résultats obtenus par Larsen et ses collègues (2009) qui se sont uniquement intéressés aux réactions face à des gains ou des pertes à des jeux d'argent (cf. chapitre 2, section 1.2.).

1. Méthode

1.1. Design et stimuli

Cette étude expérimentale adopte le design mixte suivant (figure 18) :

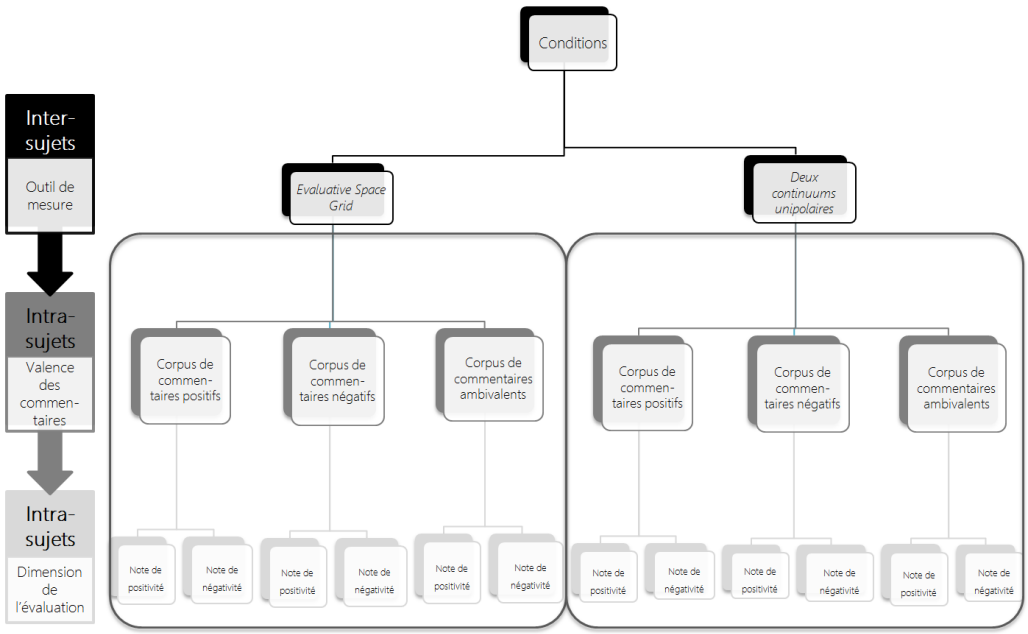
- L'outil utilisé pour mesurer la perception des commentaires par les répondants a été manipulé en mode inter-sujets : environ la moitié des répondants ont effectué leurs évaluations sur des grilles de Larsen et l'autre moitié sur des paires de continuums unipolaires.

- La valence des commentaires était manipulée en mode intra-sujets, c'est-à-dire que tous les répondants étaient exposés à trois corpus de commentaires : un corpus positif, un corpus négatif et un corpus ambivalent. Plus précisément, comme l'indique le tableau 9, il a été demandé aux participants d'évaluer deux restaurants et un hôtel suite à la lecture d'avis de consommateurs en ligne inventés pour les besoins de l'étude.

Variable « valence des commentaires »		
Corpus positif (V_{POS})	Corpus négatif (V_{NEG})	Corpus ambivalent (V_{AMB})
Restaurant le Nautilus Corpus de 4 commentaires de valence positive	Restaurant le Grand Large Corpus de 4 commentaires de valence négative	Hôtel l'Atlantique Corpus de 4 commentaires ambivalents

Tableau 9. Description de la variable « valence des commentaires »

Les répondants devaient se faire un avis sur le restaurant « le Nautilus » accompagné d'un corpus de quatre commentaires de valence positive (V_{POS}), sur le restaurant le « Grand Large » accompagné d'un corpus de quatre commentaires de valence négative (V_{NEG}) et sur l'hôtel « L'Atlantique », accompagné d'un corpus composé de quatre commentaires à la fois positifs et négatifs (V_{AMB}). Les stimuli utilisés sont détaillés dans le questionnaire présenté en annexe 1.



Tous les répondants ont évalué les 3 corpus de commentaires en termes de positivité et de négativité, attribuant 6 notes au total

Figure 18. Design de l'étude 1

- Suivant l'approche bidimensionnelle de l'attitude, chaque évaluation était décomposée en une note de positivité et une note de négativité. Dans la *condition deux continuums unipolaires*, il s'agissait des notes attribuées sur chacun des continuums, tandis que dans la *condition Evaluative Space Grid*, ces deux notes correspondaient aux coordonnées de la cellule choisie par le répondant sur l'axe de la positivité et sur l'axe de la négativité. Nous avons qualifié cette variable de « continuum de mesure » : Dans les deux conditions expérimentales, les répondants attribuaient une note sur le continuum de la positivité (C_{POS}) et une note sur le continuum de la négativité (C_{NEG}) pour chaque évaluation. Cette variable a donc été collectée en mode intra-sujets.

1.2. Participants

Grâce à un effet « boule de neige » à partir des bases de données dont nous disposions, 71 personnes ont été contactées pour participer, sur la base du volontariat, à une enquête en ligne en cliquant sur un lien reçu par mail. Les répondants étaient répartis aléatoirement dans l'une des deux conditions expérimentales : 38 en *condition Evaluative Space Grid* (ESG) et 33 en *condition deux continuums unipolaires* (2CU). Une analyse descriptive sur ce premier échantillon a révéler une différence d'âge élevée entre les deux groupes (M_{ESG} = 38,2 vs. M_{2CU} = 29,2 ; $p < 0,05$). Or, dans la mesure où cette étude porte sur des avis de consommateurs en ligne, il nous a semblé important de veiller à ce que notre échantillon soit relativement jeune, moyen indirect de garantir l'aisance des répondants avec la navigation sur le web et le commerce en ligne. Nous avons donc éliminés les répondants de plus de 40 ans, ramenant l'échantillon effectif à 59 personnes, soit 29 répondants dans la *condition Evaluative Space Grid* et 30 dans la *condition deux continuums unipolaires*. Cette procédure nous a permis de maintenir l'homogénéité des deux groupes expérimentaux en terme d'âge (M_{ESG} = 28,7 ans vs. M_{2CU} = 26,2 ans ; $p = 0,11$) et de sexe (PR_{ESG} = 61,5 % femmes vs. PR_{2CU} = 63,3 % femmes ; $p = 0,59$)[7].

1.3. Procédure

L'étude commençait par une page d'introduction correspondant à une brève mise en contexte qui invitait les répondants à se projeter dans une situation de préparation d'un weekend en amoureux au bord de la mer. Ce scenario permettait de placer les sujets dans une perspective d'évaluation de lieux visant à se restaurer et à dormir durant le séjour. C'est ainsi que la tâche de lecture des commentaires de consommateurs concernant un hôtel et deux restaurants était justifiée. Les pages de commentaires associées à chaque service apparaissaient ensuite dans un ordre aléatoire. Après avoir lu chaque corpus de commentaires, les participants devaient répondre à la question suivante : « *Pouvez-vous maintenant nous donner l'opinion (positive et négative) que vous avez de ce restaurant (hôtel)* ? ». L'annexe 1 présente en détail le questionnaire utilisé.

[7] *Nous devons toutefois relativiser ces données sociodémographiques dans la mesure où, suite à un problème informatique sur la dernière page du questionnaire, ces informations n'ont pas été enregistrées pour 16 participants de la condition Evaluative Space Grid. Cela explique sans doute la significativité marginale du test de différence de moyennes concernant l'âge des participants entre les deux conditions, même après élimination des répondants de plus de 40 ans.*

Dans la *condition Evaluative Space Grid*, une présentation de l'outil précédait la première page de commentaires (annexe 1). L'évaluation s'effectuait ensuite sur une grille de Larsen 5 x 5 similaire à celle de la figure 19. Les axes mesurant les opinions négatives et positives étaient gradués de « pas du tout » à « extrêmement ».

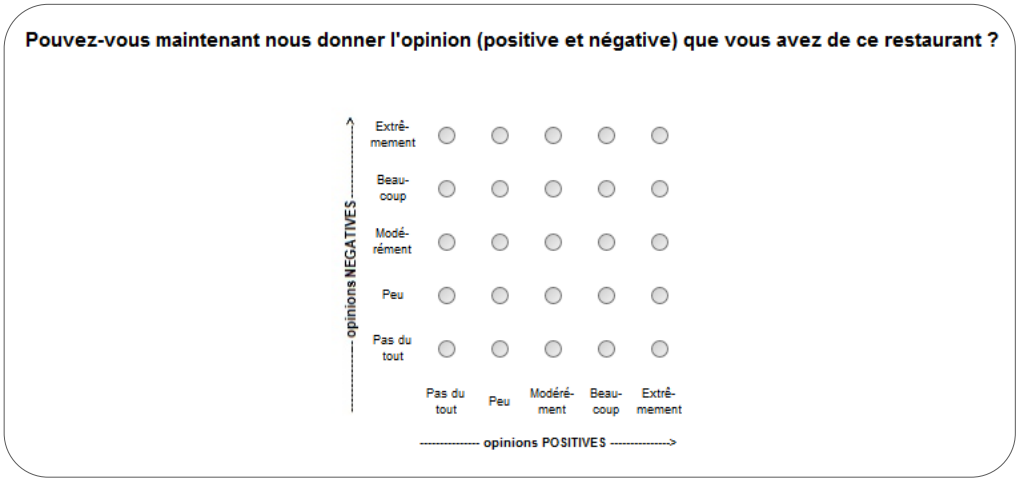


Figure 19. Item et outil de mesure utilisés pour évaluer les corpus de commentaires de consommateurs dans la *condition Evaluative Space Grid*

Les répondants affectés à la *condition deux continuums unipolaires* évaluaient l'hôtel et les deux restaurants sur des paires de deux continuums gradués de « pas du tout » à « extrêmement » et introduits par les instructions proposées par Kaplan (1972) (figure 20).

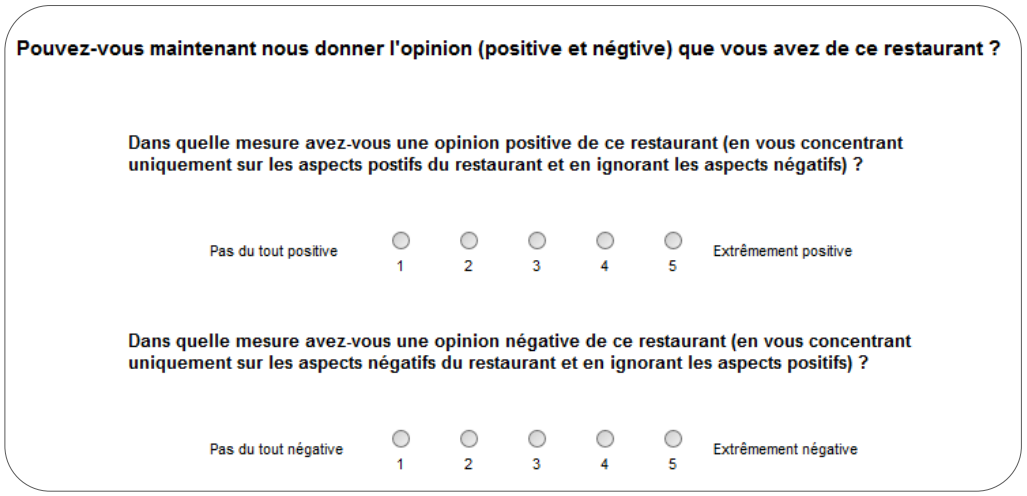


Figure 20. Items et outils de mesure utilisés pour évaluer les corpus de commentaires de consommateurs dans la *condition deux continuums unipolaires*

2. Analyses

Pour évaluer la validité convergente de la grille, nous avons tout d'abord comparé les évaluations des corpus de commentaires positifs, négatifs et ambivalents dans les deux conditions de mesure (*Evaluative Space Grids* et *deux continuums unipolaires*). Pour effectuer cette comparaison, une procédure de recodage préalable des données obtenues avec la grille de Larsen a été nécessaire. En effet, cet outil nous a fourni pour chaque corpus évalué une note allant de 1 à 25 (puisque nous avons utilisé une matrice 5 x 5) correspondant au croisement entre la note attribuée sur le continuum de la positivité et sur celui de la négativité (figure 21).

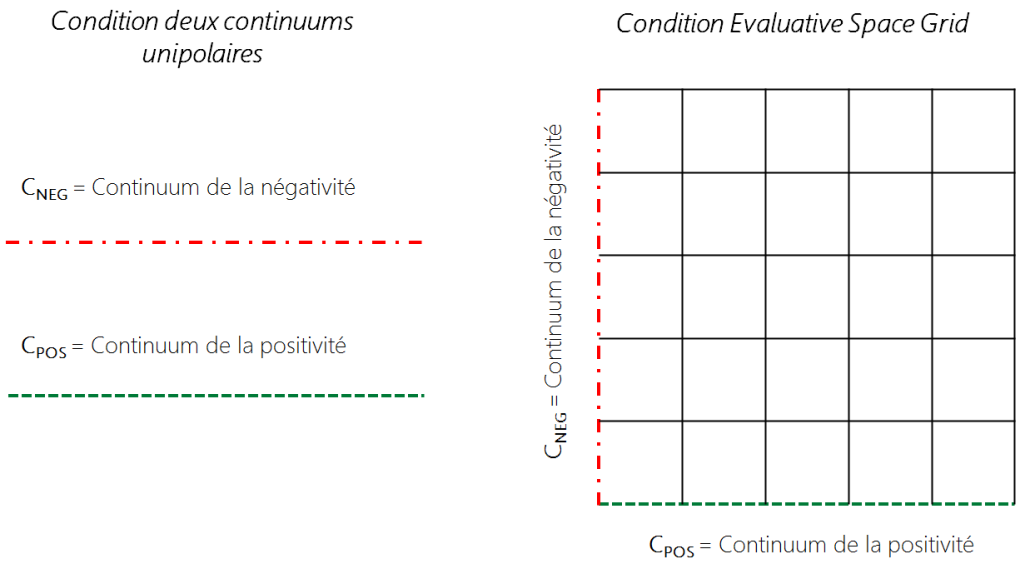


Figure 21. Format des données dans les deux conditions pour la variable « continuum de mesure »

Dans le but de comparer pour chaque stimulus la note obtenue avec la grille de Larsen aux deux notes issues des continuums unipolaires, il a été nécessaire de transformer la note issue de la grille de Larsen en deux notes : une pour les opinions positives et une pour les opinions négatives (figure 21).

Chaque point de la grille a été recodé avec deux notes correspondant à ses coordonnées de positivité et de négativité comme cela est présenté dans la figure 22.

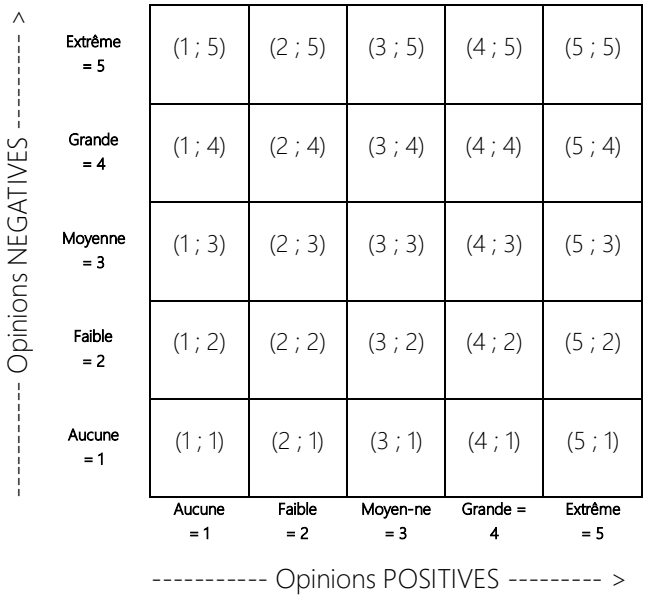


Figure 22. Codage de l'*Evaluative Space Grid* utilisé dans l'étude 1

3. Résultats

Dans le but de mettre en évidence la similarité des évaluations effectuées avec les deux outils de mesure (ESG, l'*Evaluative Space Grid* ou 2CU, les paires de deux continuums unipolaires), la variable dépendante correspondant à l'évaluation des commentaires et composée de deux notes (C_{POS}, la note de positivité et C_{NEG} la note de négativité) a été soumise à un test de différence de moyenne. Plus précisément, nous avons effectué une ANOVA mixte au design suivant : 2 (outil de mesure : ESG ou 2CU) x 3 (valence des commentaires : V_{POS}, V_{NEG} et V_{AMB}) x 2 (Continuum de mesure : C_{POS} et C_{NEG}). Dans cette analyse de la variance, la variable « outil de mesure » a été traitée en inter-sujets, tandis que les variables « valence des commentaires » et « continuum de mesure » ont été traitées en intra-sujets.

La première partie de l'ANOVA visait à vérifier que de la valence des commentaires avait bien été perçue. Cette ANOVA a révélé un effet principal significatif de la valence des corpus ($F(2;114) = 12,39$; $p < 0,01$; $\eta_p^2 = 0,18$) : le contenu des commentaires a impacté significativement les évaluations, c'est-à-dire que les notes données sur le continuum de la positivité et sur le continuum de la négativité étaient différentes en fonction de la valence des corpus de commentaires évalués. L'ANOVA a également montré un effet principal marginalement significatif du continuum de mesure de l'évaluation ($F(1 ; 57) = 3,756$; $p = 0,058$; $\eta_p^2 = 0,062$). En effet, la comparaison des moyennes des trois notes données (pour

le corpus positif, pour le corpus négatif et pour le corpus ambivalent) sur le continuum de la positivité et des trois notes données sur le continuum de la négativité, a révélé une différence marginalement significative entre les notes données sur le continuum de la positivité et les notes données sur le continuum de la négativité. Ainsi, la moyenne des notes données aux trois corpus lisse l'effet du continuum de mesure : il n'est pas surprenant que la note moyenne de positivité pour les corpus positif, négatif et ambivalent soit proche de la note moyenne de négativité pour ces trois corpus. Enfin, l'analyse a mis en évidence un effet d'interaction « valence du commentaire x continuum de mesure » ($F(2 ; 114) = 328,92$; $p < 0,01$; $\eta_p^2 = 0,85$). Cet effet correspond à l'inversion des scores entre le continuum de la positivité et le continuum de la négativité en fonction de la valence du commentaire : le corpus positif a obtenu un score plus élevé sur le continuum de la positivité que sur le continuum de la négativité alors que le corpus négatif a obtenu un score plus élevé sur le continuum de la négativité que sur le continuum de la positivité.

Par ailleurs, les analyses n'ont pas révélé d'effet d'interaction « outil de mesure x valence du corpus x continuum », ($p = 0,37$). Cela indique que, pour chaque corpus de commentaires (V_{POS}, $_{VNEG}$, V_{AMB}), les scores obtenus sur le continuum de mesure de la positivité et de la négativité étaient identiques avec les deux outils de mesure (*Evaluative Space Grid* et deux continuums unipolaires) (tableau 10).

	Ddl	F	Sig.	η_p^2
Valence stimulus	2	12,39	0,000	0,18
Valence stimulus * Outil de mesure	2	1,50	0,228	0,03
Erreur(Valence Stimulus)	114			
Evaluation	1	3,76	0,058	0,06
Evaluation * Outil de mesure	1	0,99	0,325	0,02
Erreur(Evaluation)	57			
Valence stimulus * Evaluation	2	328,92	0,000	0,85
Valence stimulus * Evaluation * Outil de mesure	2	0,99	0,373	0,02
Erreur(Valence stimulus*Evaluation)	114			

Ddl : degrés de liberté ; F : Statistique de Fisher ; Sig. : significativité ; η_p^2 : Eta partiel au carré

Tableau 10. Résultats de l'ANOVA mixte : 2 (outil de mesure) x 3 (valence des commentaires) x 2 (continuum de mesure)

Pour mieux comprendre la nature des effets, nous avons exploré le sens de l'interaction « valence du corpus x continuum de mesure » à l'aide d'une comparaison de moyennes entre les évaluations obtenues avec les deux outils. Sans surprise, quel que soit l'outil de mesure utilisé :

- Sur le continuum de la positivité, le corpus positif a été évalué plus positivement (M = 4,29) que le corpus négatif (M = 1,61) ;
- Sur le continuum de la négativité, le corpus négatif a été évalué plus négativement (M = 4,71) que le corpus positif (M = 1,76) ;

Résultat plus important pour notre propos, la mesure du corpus de commentaires ambivalents est valide puisque, quel que soit l'outil de mesure utilisé :

- Sur le continuum de la positivité, le corpus ambivalent a été évalué plus positivement (M = 3,46) que le corpus négatif (M = 1,61) ;
- Sur le continuum de la négativité, le corpus ambivalent a été évalué plus négativement (M = 3,36) que le corpus positif (M = 1,76) ;

Ces résultats sont représentés graphiquement en figure 23.

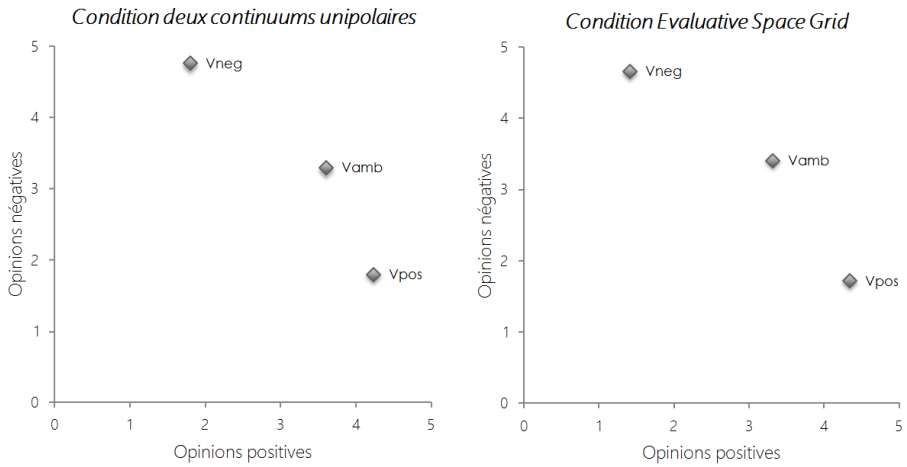


Figure 23. Scores moyens sur les continuums de la positivité et de la négativité pour les trois corpus de commentaires selon l'outil de mesure utilisé

Nous avons ensuite exploré une piste d'analyse suggérée dans les travaux de Russell et Carroll (1999). Selon ces auteurs, une des limites des mesures utilisant deux continuums

unipolaires serait qu'elles créeraient une « ambivalence artificielle », en raison de la présence des deux items. En effet, la nécessité de répondre à deux questions inciterait les répondants à rapporter des réactions négatives même en cas d'exposition à des stimuli purement positifs ou purement négatifs. Si l'on suit ce raisonnement, la note de positivité attribuée au corpus négatif devrait être plus élevée en *condition deux continuums unipolaires* qu'en *condition Evaluative Space Grid*. De même, la note de négativité attribuée au corpus positif devrait être plus élevée en *condition deux continuums unipolaires* qu'en *condition Evaluative Space Grid*. Si ces hypothèses sont vérifiées, cela apporterait un premier élément de preuve indiquant la supériorité de la grille de Larsen sur les deux continuums unipolaires.

Pour les tester, nous avons réalisé une ANOVA à un facteur (outil de mesure : ESG ou 2CU) dans le but comparer les moyennes obtenues avec chacun des outils pour le corpus positif sur le continuum négatif d'une part, et pour le corpus négatif sur le continuum positif, d'autre part. Le résultat du test omnibus initial n'était pas significatif en ce qui concerne la note de négativité attribuée au corpus positif (p = 0,725). Il n'y a donc pas eu de différence significative d'évaluation entre les deux outils.

		Somme des carrés	Ddl	Moyenne des carrés	F	Sig.	η_p^2
Corpus positif – note de négativité	Inter-groupes	,085	1	,085	,125	,725	,002
	Intra-groupes	38,593	57	,677			
	Total	38,678	58				
Corpus négatif – note de positivité	Inter-groupes	2,199	1	2,199	3,705	,059	,061
	Intra-groupes	33,834	57	,594			
	Total	36,034	58				

Ddl : degrés de liberté ; F : Statistique de Fisher ; Sig. : significativité ; η_p^2 : Eta partiel au carré

Tableau 11. Comparaisons de moyennes entre les conditions pour la note de négativité attribuée au corpus positif et pour la note de positivité attribuée au corpus négatif

Par contre, comme nous l'avions postulé, le résultat est marginalement significatif en ce qui concerne la note de positivité attribuée au corpus négatif, indiquant que les notes de positivité ont eu tendance à être artificiellement plus élevées en réaction à un stimulus négatif lorsque la mesure était effectuée avec deux continuums unipolaires plutôt qu'avec une grille de Larsen ($F(1 ; 57) = 3,705$; $p = 0,059$; $\eta_p^2 = 0,061$). Ainsi, l'hypothèse de

« l'ambivalence artificielle » créée par les deux continuums unipolaires est vérifiée pour la mesure de la positivité. Ces résultats sont présentés dans le tableau 11.

4. Conclusion

L'étude 1 apporte la preuve initiale de la validité convergente de la grille de Larsen en comparant les évaluations issues de cet outil aux évaluations effectuées avec des paires de deux continuums unipolaires. En effet, au global, quel que soit l'outil de mesure utilisé, les notes attribuées sur le continuum de la positivité et sur le continuum de la négativité n'étaient pas significativement différentes pour les trois types de stimuli : le corpus de commentaires positifs, le corpus de commentaires négatifs et le corpus de commentaires ambivalents (figure 23).

Par contre, lorsque l'on a comparé plus spécifiquement la note de positivité attribuée au corpus négatif en fonction de l'outil de mesure utilisé, celle-ci était significativement plus faible en *condition Evaluative Space Grid*. Ce résultat fournit un début de preuve de la supériorité de la grille de Larsen sur les mesures utilisant deux continuums unipolaires. En effet, ces dernières sont associées dans la littérature à des évaluations « artificiellement ambivalentes » dans la mesure où la présence de deux continuums peut inciter les répondants à rapporter des réactions allant dans les deux directions évaluatives alors que leur opinion réelle est polarisée (Russell et Carroll, 1999). Au contraire, la grille de Larsen semble limiter certaines erreurs de réponse.

Cette réplication partielle des résultats de Larsen et ses collègues (2009) appliquée à la mesure des opinions positives et négatives suite à la lecture d'avis de consommateurs en ligne, nous encourage à explorer davantage l'intérêt de la grille de Larsen.

5. Limites

Cette étude n'est cependant pas exempte de limites. Pour commencer, le nombre de stimuli était très restreint (un stimulus positif, un stimulus négatif et un stimulus ambivalent), ce qui limite le potentiel de généralisation des résultats. Cela peut peut-être expliquer pourquoi nous n'avons pas réussi, pour le corpus négatif, à montrer que la grille de Larsen offre une meilleure mesure des stimuli univalents en réduisant le phénomène d'ambivalence artificielle associée aux deux continuums unipolaires dans la littérature. Il apparait donc nécessaire de répliquer les résultats.

Ensuite, contrairement au protocole suivi par Larsen et ses collègues (2009), le logiciel de collecte de données que nous avons utilisé ne nous permettait pas d'effectuer une rotation aléatoire des continuums de mesure des opinions positives et négatives dans la *condition deux continuums unipolaires*. Nous pensons toutefois que cette limite est mineure dans la mesure où il aurait pu sembler étonnant d'effectuer une telle rotation aléatoire des questions dans la *condition deux continuums unipolaires* alors que dans la condition Evaluative Space Grid l'orientation des axes ne variait pas (opinions négatives à la verticale et opinions positives à l'horizontale). Rappelons à cet égard que Larsen et ses collègues (2009), ainsi que Andrade et Cohen (2007) ont montré que l'orientation des axes de la grille n'a aucun impact sur les résultats obtenus. Cela nous incite à penser que l'ordre d'apparition des continuums unipolaires n'a pas eu d'effet sur les réponses en *condition deux continuums unipolaires*.

Par ailleurs, le mode de recueil des données grâce à un effet « boule de neige », peut soulever certains problèmes dans le cas d'un design expérimental. En effet, dans la mesure où nous avons demandé aux personnes contactées de faire suivre le lien de l'étude à leurs contacts, il était difficile de garantir l'homogénéité des échantillons.

Enfin, un élément important n'a pas été exploré dans le cadre de cette étude : la performance des différents outils dans le cadre de la mesure de l'indifférence. En effet, le domaine d'application des commentaires de consommateurs se prêtait mal à la production de stimuli indifférents puisque le fait même de donner spontanément son avis sur quelque chose est contraire à la notion d'indifférence. Nous étudierons donc comment la grille permet de capturer distinctement les stimuli indifférents et ambivalents dans la suite de la partie empirique de la thèse.

Section 2. Analyse de la validité convergente de la grille avec les différentiels sémantiques, les discours des consommateurs et une mesure de l'ambivalence subjective (Etude 2)

Cette deuxième étude vise à prolonger notre travail préalable d'analyse de la validité convergente de l'*Evaluative Space Grid* en la comparant cette fois-ci plus précisément avec des dispositifs qu'elle pourrait concurrencer. Nous soutenons que la grille de Larsen permet de saisir la part d'ambivalence contenue dans les évaluations globales de consommateurs. Appliquée à la mesure de la satisfaction, nous avons donc cherché à nous assurer, d'une part qu'elle convergeait avec d'autres mesures de la satisfaction et d'autre part, qu'elle convergeait avec d'autres mesures de l'ambivalence. Pour ce faire, nous avons comparé les scores de satisfaction globale obtenus avec trois outils différents : des *Evaluative Space Grids*, des différentiels sémantiques et des verbatims de consommateurs décrivant les éléments contribuant à leur satisfaction et à leur insatisfaction. D'autre part, nous avons comparé les notes issues de la grille de Larsen à une mesure de l'ambivalence subjective.

1. Méthode

1.1. Design et stimuli

Dans cette étude les participants devaient évaluer leur satisfaction à l'égard de deux services : les restaurants McDonald's (première partie de l'enquête) et leur dernière expérience chez le coiffeur (deuxième partie de l'enquête). L'accès limité aux répondants, via une collecte de données en ligne auprès de nos contacts personnels, a orienté le choix de ces objets d'évaluation. En effet, s'intéresser à la satisfaction suppose d'interroger des individus sur des services qu'ils ont effectivement consommés. McDonald's étant la plus grande chaine de restauration rapide au monde, il est fortement probable qu'une personne interrogée au hasard en soit client ou l'ait été. De même, la plupart des gens vont chez le coiffeur. Ces deux services constituent donc des domaines d'application facilitant l'accès aux répondants.

Dans le cadre d'un design intra-sujet, les participants ont été invités à évaluer ces services avec les quatre outils suivants : une *Evaluative Space Grid*, un différentiel sémantique, deux questions ouvertes portant sur les éléments ayant contribué à leur satisfaction et leur

insatisfaction et une mesure subjective de l'ambivalence. Tous les participants ont donc effectué l'évaluation des deux services avec les quatre outils de mesure.

1.2. Participants

Grâce à un effet « boule de neige » à partir des bases de données dont nous disposions, 130 personnes ont participé, sur la base du volontariat, à une enquête en ligne en cliquant sur un lien reçu par mail. Les répondants étaient âgés de 23 à 68 ans (âge moyen : 35 ans) et 65% d'entre eux étaient des femmes. Suite à un problème informatique, les réponses aux questions ouvertes portant sur les motifs de satisfaction et d'insatisfaction des 5 premiers répondants n'ont pas été enregistrées, ramenant notre échantillon à 125 personnes. Les données ont ensuite été épurées séparément pour chaque service évalué. Ainsi, dans l'ensemble de données associé à l'appréciation du McDonald's, 6 répondants ont été supprimés parce qu'ils n'ont pas répondu aux questions ouvertes (N_{McDo} = 119 personnes), et 11 autres pour la même raison dans la base de données portant sur l'évaluation du coiffeur. 2 répondants ont également été supprimés de l'analyse car la lecture de leur réponse aux questions ouvertes indiquait qu'ils n'avaient pas bien compris ce qui leur était demandé. Suite à cette procédure d'épuration, l'échantillon effectif était de 112 personnes. La fatigue des répondants lors de la passation de la deuxième partie du questionnaire peut expliquer pourquoi nous avons dû épurer davantage la base de données associée à l'évaluation du coiffeur.

1.3. Procédure

L'étude commençait par une page d'introduction expliquant aux répondants les tâches d'évaluation à effectuer. Les pages qui suivaient visaient à familiariser les participants avec la grille de Larsen. Après une brève explication du fonctionnement de cet outil, trois exemples d'évaluation de commentaires de consommateurs permettaient aux répondants de s'entrainer. Une correction était proposée à la suite de chaque exemple dans le but de rassurer les participants sur leur compréhension de l'utilisation de l'outil.

A l'issue de cette phase préliminaire, l'évaluation de chaque service s'effectuait de manière graduelle. En d'autres termes, les questions allaient du général au particulier, dans l'ordre suivant :

1) La satisfaction globale était tout d'abord évaluée sur un différentiel sémantique en 9 points (allant « d'extrêmement insatisfait » à « extrêmement satisfait ») avec la question suivante : « *Quel est votre niveau de satisfaction vis-à-vis de vos expériences au McDonald's (resp. votre dernière expérience chez le coiffeur) ?* ».

2) Puis, les répondants étaient invités à préciser leur opinion sur une grille mesurant le niveau de satisfaction en abscisse et le niveau d'insatisfaction en ordonnée (les axes étaient gradués de « aucune » à « extrême »). La question « *Plus précisément, quel est votre niveau de satisfaction / insatisfaction vis-à-vis vos expériences au McDonald's (resp. votre dernière expérience chez le coiffeur) ?* » permettait d'introduire cette deuxième étape.

3) Il était ensuite demandé aux participants de motiver leur réponse dans des espaces de réponses libres permettant de décrire les « éléments ayant contribué à leur satisfaction » et les « éléments ayant contribué à leur insatisfaction ».

4) Cette procédure d'évaluation se terminait par la mesure du sentiment subjectif d'ambivalence, effectuée à l'aide d'une adaptation de l'échelle proposée par Conner et Spark (2002) : « *J'ai des opinions contradictoires au sujet de…* », « *J'ai un avis partagé vis-à-vis de…* », « *Je suis indécis(s) au sujet de…* » et « *J'ai des opinions positives et négatives au sujet de…* ». Les quatre items étaient mesurés avec des échelles de Likert en 7 points (allant de « pas du tout d'accord » à « tout à fait d'accord »).

Une question concernant la fréquence d'utilisation des services évalués permettait ensuite de nous assurer que tous les répondants étaient bien clients du McDonald's et qu'ils se rendaient de temps en temps chez le coiffeur. Aucune fréquence de consommation nulle n'ayant été rapportée, il est possible que les personnes concernées aient abandonné le remplissage du questionnaire en cours de route. La dernière page du questionnaire visait à collecter des informations sociodémographiques sur les participants. L'intégralité du questionnaire est présentée en annexe 2.

2. Analyses

L'analyse suppose de comparer les évaluations issues de la grille de Larsen aux évaluations effectuées avec les trois autres outils : les verbatims récoltés à l'aide des deux questions

ouvertes, une mesure de la satisfaction provenant d'un différentiel sémantique et une mesure de l'ambivalence subjective. Cette section présente la préparation des données effectuée dans le but de rendre comparables les informations collectées avec les différents outils (2.1) et les analyses réalisées dans le cadre de ces comparaisons (2.2).

2.1. Préparation des données

Nous décrivons dans cette partie les procédures de recodage mises en œuvre pour rendre comparables les données issues des différents outils.

En vue de la comparaison des notes issues de la grille au discours des consommateurs

Comme dans l'étude 1, chaque point de l'*Evaluative Space Grid* a été recodé par deux notes correspondant à ses coordonnées sur l'axe de la satisfaction et sur l'axe de l'insatisfaction (cf. figure 22 en page 85). Les deux notes obtenues prenaient des valeurs allant de 1 pour « aucune satisfaction (resp. insatisfaction) » à 5 pour « extrême satisfaction (resp. insatisfaction) ».

Les verbatims issus des deux questions ouvertes invitant les répondants à expliciter les « éléments ayant contribué à leur satisfaction » et les « éléments ayant contribué à leur insatisfaction » ont été codés de 1 à 5 de manière à distinguer 5 degrés de satisfaction et 5 degrés d'insatisfaction. Chaque idée différente s'est vue attribuée un point dans la limite de 4 points, ce qui nous a permis de prendre en compte les situations dans lesquelles les répondants n'ont rien répondu, codées 0. Nous avons ensuite ajouté 1 point partout pour obtenir le même codage que pour les axes de la grille de Larsen (allant de 1 = « aucune satisfaction (insatisfaction) » à 5 = « extrême satisfaction (insatisfaction) ».

	McDonald's		Coiffeur	
	Verbatims	ESG	Verbatims	ESG
Satisfaction	SAT_{VERB}	SAT_{ESG}	SAT_{VERB}	SAT_{ESG}
Insatisfaction	$INSAT_{VERB}$	$INSAT_{ESG}$	$INSAT_{VERB}$	$INSAT_{ESG}$

Tableau 12. Comparaison entre les verbatims recodés et les notes issues de la grille de Larsen

Nous avons ainsi obtenu des paires de deux notes pouvant être comparées suivant l'outil de mesure dont elles étaient issues : la grille (SAT_{ESG} ; $INSAT_{ESG}$) et les verbatims (SAT_{VERB} ; $INSAT_{VERB}$) (Tableau 12).

En vue de la comparaison des notes issues de la grille à celles des différentiels sémantiques

Dans cette étude, les évaluations effectuées avec des différentiels sémantiques nous ont fourni des notes allant de 1 à 9 (avec 1 = « extrême insatisfaction » et 9 = « extrême satisfaction »). Or, effectuer une évaluation sur un différentiel sémantique suppose pour le répondant d'effectuer dans sa tête le solde net entre son niveau de satisfaction et son niveau d'insatisfaction. Suivant cette logique, pour rendre comparables les notes issues de la grille à celles provenant des différentiels sémantiques, nous avons appliqué la formule suivante :

$$Transformation\ ESG\ en\ format\ DS = SAT_{ESG} - INSAT_{ESG} + 5$$

En effet, il a été nécessaire d'ajouter 5 au solde net des notes de satisfaction et d'insatisfaction issues de la grille pour qu'elles prennent des valeurs allant de 1 à 9 (et non de -4 à 4). Nous avons ainsi obtenu à partir de la grille un index bipolaire de la satisfaction.

En vue de la comparaison entre les notes de la grille et le score d'ambivalence subjective

Pour chaque répondant nous avons calculé un score d'ambivalence en effectuant la moyenne des notes données aux quatre items de l'échelle d'ambivalence subjective. Ce score d'ambivalence prend des valeurs allant de 1 à 7 puisque les items avaient été évalués sur des échelles de Likert en 7 points. Nous souhaitions comparer ce score à une note d'ambivalence objective calculée à partir de la grille. Pour ce faire, nous avons utilisé les coordonnées des cellules de la grille (SAT_{ESG} et $INSAT_{ESG}$) pour calculer un score d'ambivalence objective avec la formule de Thompson et ses collègues (1995). Toutefois, l'index proposé dans cet article de référence évoluant de -1 à 5, nous avons ajouté 2 à la formule initiale pour obtenir des notes allant de 1 à 7. Nous avons donc appliqué la formule suivante en utilisant les notes issues de la grille (codage obtenu présenté en figure 24) :

$$Ambivalence\ objective = [\ (SAT_{ESG} + INSAT_{ESG})/2 - |\ SAT_{ESG} - INSAT_{ESG}\ |] + 2$$

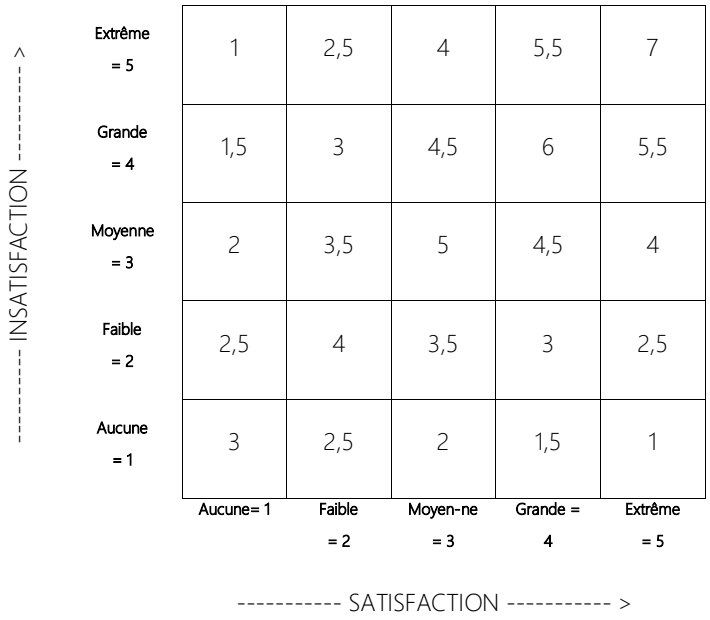


Figure 24. Application de la formule de Thompson et al. (1995) aux notes issues de la grille de Larsen

2.2. Tests statistiques effectués

Nous avons effectué deux types d'analyses pour comparer les notes obtenues avec ces quatre outils : (1) des corrélations pour voir s'il existe une relation de dépendance linéaire entre les données issues de la grille et des autres outils et (2) des tests de différence de moyennes pour voir si les notes obtenues diffèrent significativement en fonction de l'outil de mesure utilisé.

3. Résultats

3.1. Comparaison des évaluations issues de la grille avec le discours évaluatif des consommateurs

Nous avons tout d'abord calculé des corrélations de Pearson entre les coordonnées des notes de satisfaction et d'insatisfaction issues de la grille et les notes issues des verbatims de satisfaction et d'insatisfaction recodés concernant l'évaluation des expériences au McDonald's. La même analyse est conduite pour les évaluations des participants concernant leur dernière expérience chez le coiffeur. Les résultats indiquent qu'il existe une relation linéaire significative entre les notes issues des deux outils (tableau 13). Ces

corrélations de Pearson sont modérées sauf dans le cas des notes d'insatisfaction concernant le McDonald's où la relation est relativement faible.

	McDonald's	Coiffeur
	Corrélation verbatims / ESG	Corrélation verbatims / ESG
Satisfaction	0,54 ($p < 0,001$)	0,48 ($p < 0,001$)
Insatisfaction	0,28 ($p < 0,001$)	0,50 ($p < 0,001$)

Tableau 13. Résultats des corrélations entre les verbatims recodés et les notes issues de la grille de Larsen

Pour approfondir cette analyse, nous avons effectué des tests de différence de moyennes sur mesure répétées entre les notes issues des deux outils. Les résultats indiquent une différence de moyenne significative entre les notes issues de la grille et des verbatims pour la mesure de l'insatisfaction concernant le McDonald's et pour la mesure de la satisfaction concernant la dernière expérience chez le coiffeur (tableau 14). Ces conclusions sont cohérentes avec la première phase d'analyse puisqu'il s'agissait des évaluations pour lesquelles les corrélations étaient les plus faibles. Pour comprendre ces résultats, nous avons analysé le sens des différences de moyennes. Pour l'évaluation du McDonald's, les verbatims d'insatisfaction obtiennent des notes supérieures à celles issues de la grille ($INSAT_{ESG} < INSAT_{VERB}$), tandis que c'est le contraire pour les notes de satisfaction en ce qui concerne la dernière expérience chez le coiffeur ($SAT_{ESG} > SAT_{VERB}$).

Ces résultats reflètent probablement une difficulté à prendre en compte, dans notre codage, l'intensité ou l'importance de certains verbatims. En effet, on peut penser que le fait d'être satisfait de « la coupe », (verbatim codé 1) a pu se traduire par une évaluation de satisfaction « grande » ou « extrême », soit une note de 4 ou 5 sur la grille. Au contraire, lorsqu'il s'agissait d'évaluer le McDonald's, certains répondants se sont montrés particulièrement prolixes pour décrire tous les griefs qu'ils souhaitaient adresser au géant de la restauration rapide, ce qui a impliqué fréquemment des notes de 5 lors du codage des verbatims d'insatisfaction. Il est légitime de penser que ces discours, parfois politisés sur le McDonald's, ne se seraient pas forcément traduit par des notes d'insatisfaction si élevées s'ils avaient été rapportés sur la grille de Larsen.

	McDonald's		Coiffeur	
	Verbatims	ESG	Verbatims	ESG
Satisfaction	M_{VERB} = 2,09	M_{ESG} = 2,01	M_{VERB} =2,10	M_{ESG} =2,67
	p = 0,390 ; η_p^2 = 0,006		p < 0,001 ; η_p^2 = 0,258	
Insatisfaction	M_{VERB} = 2,81	M_{ESG} = 2,15	M_{VERB} = 1,28	M_{ESG} = 1,44
	p < 0,001 ; η_p^2 = 0,264		p = 0,083 ; η_p^2 = 0,027	

Tableau 14. Résultats des ANOVA sur mesures répétées entre les verbatims recodés et les notes issues de la grille de Larsen

Cette première série de comparaison va dans le sens de la validité convergente de l'*Evaluative Space Grid* avec les discours réels des consommateurs. Elle soulève toutefois un éventuel décalage entre la quantité des propos, c'est-à-dire le nombre d'attributs codés dans les verbatims et le poids accordé à chaque attribut. On retrouve ici l'idée d'importance relative des attributs dans la composition de la satisfaction décrite dans les différents modèles de formation de l'attitude (Einhorn, 1970 ; Fishbein, 1963 ; 1967 ; Green et Wind, 1973 ; Lehmann, 1971 ; Wright, 1973). En effet, la décomposition de l'évaluation globale en deux notes dans le cadre d'une évaluation effectuée sur la grille facilite probablement la mise en œuvre de ces modèles décisionnels pour les répondants.

3.2. Comparaison des évaluations issues de la grille de Larsen avec les notes obtenues sur les différentiels sémantiques

Les corrélations entre les notes issues de la grille de Larsen (différence entre la note de satisfaction et d'insatisfaction) et les notes obtenues sur les différentiels sémantiques sont fortes pour les deux services évalués dans le cadre de cette étude puisqu'elles sont supérieures à 0,7 pour l'évaluation du McDonald's et de la dernière expérience chez le coiffeur. Les résultats détaillés sont présentés dans le tableau 15 ci-après.

McDonald's	Coiffeur
Corrélation différentiel sémantique /ESG	Corrélation différentiel sémantique / ESG
0,86 (p < 0,001)	0,79 (p < 0,001)

Tableau 15. Résultats des corrélations entre les notes des différentiels sémantiques et de la grille de Larsen

Cette analyse a été approfondie en effectuant des tests de différences de moyennes pour mesures répétées. Le tableau 16 montre que les évaluations ne varient pas significativement selon l'outil utilisé. Notons toutefois une différence marginalement significative dans le cas du McDonald's avec un effet très faible (η_p^2 = 0,027).

McDonald's		Coiffeur	
Différentiel sémantique	ESG	Différentiel sémantique	ESG
M_{DS} = 4,70	M_{ESG} = 4,86	M_{DS} = 6,40	M_{ESG} = 6,23
p = 0,074 ; ηp^2 = 0,027		p = 0,145 ; ηp^2 = 0,019	

Tableau 16. Résultats des ANOVA sur mesures répétées entre les différentiels sémantiques et les notes issues de la grille de Larsen

Cette deuxième phase de comparaison tend à confirmer la validité convergente de la grille de Larsen avec le traditionnel différentiel sémantique.

3.3. Comparaison des évaluations issues de la grille de Larsen avec les scores moyens d'ambivalence subjective

Les corrélations entre les scores d'ambivalence objective calculés à partir des coordonnées de la grille de Larsen (avec application de la formule de Thompson et ses collègues) et les scores d'ambivalence subjective calculés à partir des 4 items de mesure du ressenti d'ambivalence de Conner et Spark (2002) sont significatives pour l'évaluation du McDonald's et du coiffeur. Elles indiquent toutes deux une relation linéaire forte entre les notes issues des deux outils, comme l'indique le tableau 17.

McDonald's		Coiffeur	
Ambivalence subj.	Ambivalence $obj._{ESG}$	Ambivalence subj.	Ambivalence $obj._{ESG}$
0,6 (p < 0,001)		0,57 (p < 0,001)	

Tableau 17. Résultats des corrélations entre le score d'ambivalence subjective et le score d'ambivalence objective calculé à partir de la grille de Larsen

Toutefois, le test de différence de moyenne sur mesures répétées indique une différence significative pour les notes issues des deux outils dans le cas de l'évaluation de la dernière expérience chez le coiffeur comme le montre le tableau 18.

Ces résultats sont conformes à ceux rapportés dans la littérature. En effet, dans les recherches de Thompson et ses collègues (1995) et de Priester et Petty (1996) le taux de corrélation entre ces deux mesures se situait aux alentours de 0,4. Cette faible relation suggère que l'expérience d'une attitude ambivalente n'implique pas nécessairement un ressenti subjectif d'ambivalence.

En effet, les mesures objectives de l'ambivalence sont basées sur l'idée que les répondants ne sont pas nécessairement conscients de leur ambivalence. Dans cette perspective, les mesures constituées de deux notes (une pour la positivité et une pour la négativité) permettent de capturer de manière indirecte des éléments d'évaluation qui ne sont pas nécessairement saillants à l'esprit des répondants lorsqu'ils expriment plus directement leur sentiment subjectif d'ambivalence (Jonas, Brömer et Diehl, 2000).

McDonald's		Coiffeur	
Ambivalence subj.	Ambivalence obj.$_{ESG}$	Ambivalence subj.	Ambivalence obj.$_{ESG}$
M_{AS} = 3,81	M_{ESG} = 3,98	M_{AS} = 2,60	M_{ESG} = 3,32
p = 0,129 ; η_p^2 = 0,019		p < 0,001 ; η_p^2 = 0,245	

Tableau 18. Résultats des différences de moyennes entre le score d'ambivalence subjective et le score d'ambivalence objective calculé à partir de la grille de Larsen

La grille de Larsen semble donc, d'une part, converger de manière tout à fait satisfaisante avec une mesure subjective de l'ambivalence et d'autre part, permettre de mesurer indirectement ou objectivement, l'ambivalence.

4. Conclusion

Cette étude nous a permis de comparer les évaluations issues de la grille de Larsen à celles de deux autres outils de mesure de la satisfaction et d'une mesure de l'ambivalence subjective. Les résultats confortent la validité convergente de l'*Evaluative Space Grid* en ce qui concerne la mesure de la satisfaction et de l'ambivalence. En effet, si la relation est loin d'être parfaite, il existe clairement un lien entre les évaluations issues de la grille de Larsen et les verbatims récoltés pour décrire la satisfaction et l'insatisfaction des consommateurs. Cette comparaison permet également de mettre en évidence la prise en compte de l'importance des attributs dans la mesure effectuée avec la grille de Larsen. Ensuite, les

évaluations issues de la grille de Larsen et de différentiels sémantiques sont très corrélées, indiquant que le solde effectué mentalement dans le cadre d'une mesure bipolaire peut être décomposé au sein d'une mesure bidimensionnelle de la satisfaction. Enfin, l'*Evaluative Space Grid* apporte également une mesure objective de l'ambivalence qui se corrèle bien (et même mieux que dans les travaux existants) avec une mesure subjective du sentiment d'ambivalence.

5. Limites

Cette étude comporte toutefois quelques limites. Tout d'abord, elle se borne à l'évaluation de deux services, ce qui peut restreindre la généralisation des résultats. Ceci est d'autant plus vrai que les conclusions diffèrent parfois qu'il s'agisse de l'évaluation des expériences au McDonald's ou de la dernière expérience chez le coiffeur. Cela soulève la nécessité de répliquer les résultats dans des contextes aussi variés que possibles.

D'autre part, le choix d'un design intra-sujets peut être discuté. En effet, on peut penser que l'organisation logique de la pensée a pu inciter les individus à noter de manière cohérente les objets évalués avec les différents outils de mesure. Toutefois, l'ordre des mesures, de la plus agrégée (le différentiel sémantique) à la plus détaillée (la mesure de l'ambivalence subjective) et l'impossibilité de revenir en arrière durant la passation de l'étude, ont permis, nous l'espérons de limiter cette faiblesse.

***

La validité convergente de l'outil avec différentes mesures alternatives étant établie, nous pouvons passer à l'étape suivante de notre travail consistant à valider l'*Evaluative Space Grid* comme mesure de la satisfaction client (Chapitre 4).

Chapitre 1	
Revue de la littérature	Le concept d'ambivalence La mesure de l'ambivalence
Etude préliminaire	La mesure des évaluations globales de consommateurs dans le e-commerce : pratiques et perception des internautes

↓

Chapitre 2	
Cadre conceptuel et hypothèses	Présentation de *l'Evaluative Space Grid* Résultats attendus

↓

Chapitre 3	
Etude 1	Validité convergente de *l'Evaluative Space Grid* avec les paires de continuums unipolaires
Etude 2	Validité convergente de *l'Evaluative Space Grid* avec les différentiels sémantiques, les discours des consommateurs et une mesure de l'ambivalence subjective

↓

Chapitre 4	
Etude 3	Test de *l'Evaluative Space Grid* comme outil de mesure de la satisfaction client

↓

Chapitre 5	
Etude 4	Etude qualitative après d'un échantillon d'experts : - qu'en pensent les chercheurs ? - qu'en pensent les professionnels ?

↓

Chapitre 6	
Etude 5	Place de l'incertitude dans les évaluations effectuées avec *l'Evaluative Space Grid*

– Chapitre 4 –

Etude de la pertinence de l'*Evaluative Space Grid* comme outil de mesure de la satisfaction client[8]

Après avoir mis en évidence dans le chapitre 3 la validité convergente de *l'Evaluative Space Grid*, le chapitre 4 a pour objectif de valider plus précisément la pertinence de cet outil en tant que mesure de la satisfaction globale. En effet, l'évaluation de la satisfaction est une question importante pour les chercheurs en marketing comme pour les praticiens, dans la mesure où la satisfaction est considérée comme un baromètre de la performance, pouvant prédire d'autres variables clés en marketing comme le profit ou la fidélité (Anderson, Fornell et Lehmann, 1994 ; Hurley et Estelami, 1998 ; Mittal, Ross et Baldasare, 1998). Compte tenu de son importance, la plupart des études consommateurs commencent par ou incluent une question sur la satisfaction globale, comme nous l'ont rappelés les professionnels des études que nous avons interviewés (cf. interviews avec des professionnels des études : chapitre 5, Section 2). Dans cette perspective, nous voulons démontrer que l'*Evaluative Space Grid* permet de capturer les quatre types de réactions évaluatives décrites dans le chapitre 2 de la thèse : satisfaction, insatisfaction, indifférence, ambivalence.

[8] *Cette étude a bénéficié du support de la Chaire AFNOR Performance des organisations de la Fondation Paris-Dauphine.*

Section 1. Présentation de l'étude et formulation des hypothèses

L'objectif de cette étude est de montrer que la grille fournit une meilleure mesure de la satisfaction que les traditionnels différentiels sémantiques. Nous nous appuyons sur un modèle théorique très courant en marketing qui établit une relation positive entre la qualité perçue, la satisfaction et les intentions comportementales (Cronin, Brady et Hult, 2000 ; Olsen, 2002 ; Taylor, 1997). Plus précisément, nous nous focalisons sur deux variables d'intentions comportementales fondamentales pour le marketing : l'intention de fidélité et l'intention de bouche-à-oreille. Dans le cadre de ce modèle, nous supposons que la satisfaction agit comme un médiateur partiel des effets de la qualité perçue sur les intentions comportementales. Nous anticipons également une relation positive entre intentions de fidélité et intentions de bouche-à-oreille (Anderson, Fornell et Lehmann, 1994 ; Cronin, Brady et Hult, 2000 ; Szymanski et Henard, 2001). Le modèle théorique proposé est présenté en figure 25.



Figure 25. Modèle théorique de la relation qualité perçue – satisfaction globale – intentions comportementales

Nous avons mis en place un design expérimental inter-sujets à deux conditions pour comparer l'ajustement des deux modèles de mesure alternatifs : dans le premier, la satisfaction est évaluée avec les traditionnels différentiels sémantiques (*condition DS*) ; dans le second elle est mesurée avec des *Evaluative Space Grids* (*condition ESG*). Le choix d'un design inter-sujets nous préserve d'un potentiel effet de contamination entre les deux outils de mesure (Mucchielli, 1999). En effet, l'organisation logique de la pensée pourrait, dans le cas d'un design intra-sujets, conduire les participants à ajuster la réponse apportée à l'outil présenté en second de manière à la rendre cohérente avec la réponse apportée à l'outil présenté en premier.

Dans le cadre de cette étude, nous avons formulé deux hypothèses.

La première porte sur la validité de construit. En effet, nous avons mis en évidence dans la revue de la littérature (Chapitre 1) un écart entre la conceptualisation théorique des réactions évaluatives, aujourd'hui largement considérées comme bidimensionnelles (Cacioppo et Berntson, 1994 ; Cacioppo & al., 2011 ; Cacioppo, Gardner et Berntson, 1997), et la mesure qui en est généralement faite dans le cadre d'approches globales en marketing. En particulier, de nombreux travaux en gestion envisagent la satisfaction comme un construit bidimensionnel (Herzberg, Mausner et Snyderman, 1959 ; Llosa, 1996 ; Ray et Gotteland, 2005). Nous pensons donc que l'*Evaluative Space Grid* devrait améliorer la validité du construit de satisfaction, d'où l'hypothèse H1 suivante :

> *H1 : Comparativement au modèle de mesure utilisant des différentiels sémantiques, le modèle de mesure utilisant des Evaluative Space Grids s'ajuste mieux au modèle théorique*

Nous avons d'autre part fait l'hypothèse d'une amélioration de la force de la relation entre la satisfaction globale et les intentions comportementales, lorsque la satisfaction est envisagée comme un concept bidimensionnel mesuré avec des *Evalautive Space Grids* plutôt que comme un concept unidimensionnel mesuré avec des différentiels sémantiques. Cela nous a amené à formuler l'hypothèse suivante :

> *H2 : Comparativement au modèle de mesure utilisant des différentiels sémantiques, le modèle de mesure utilisant des Evaluative Space Grids offre une meilleure prédiction des intentions de fidélité et de bouche-à-oreille.*

Section 2. Méthode

1. Choix de l'objet d'étude

La performance de l'*Evaluative Space Grid* a été mise à l'épreuve dans le cadre d'une étude de la satisfaction à l'égard de la banque. Comme nous l'avons expliqué précédemment, nous avons décidé de tester cet outil dans le domaine des services, secteur d'activité propice à la formation d'évaluations ambivalentes. Le secteur bancaire a été retenu au terme d'un focus group mené avec sept individus au sujet de leurs opinions sur divers

services. Il en est ressorti que tous les participants avaient une opinion sur leur banque et que ce service semblait susciter des évaluations mitigées.

2. Procédure

Les sujets étaient répartis aléatoirement en deux groupes correspondant aux deux conditions expérimentales. Les répondants affectés à la *condition ESG* ont rempli une enquête de satisfaction à l'égard de leur banque dans laquelle l'évaluation du service était mesurée par trois items de mesure des réactions globales à compléter sur des *Evaluative Space Grids*. Les répondants affectés à la *condition DS* ont rempli exactement le même questionnaire sauf que les trois items visant à capturer la satisfaction globale étaient mesurés avec des différentiels sémantiques. Il était précisé aux participants de se focaliser uniquement sur leur banque principale, c'est-à-dire celle avec laquelle ils sont le plus souvent en contact, dans le cas où ils seraient clients de plusieurs établissements. Les données ont été collectées par un institut d'études professionnel, Easy Panel, dans le cadre d'une étude administrée en ligne. Ce recueil de données a été possible grâce au support financier de la Chaire AFNOR Performance des organisations de la Fondation Paris-Dauphine.

3. Participants

400 sujets ont complété l'étude, soit 200 personnes par condition expérimentale. Nous avons effectué une première phase d'épuration des données en supprimant les réponses remplies de manière aléatoire par les panélistes. Ces réponses ont été repérées en utilisant les deux critères suivants :

- Le temps de réponse chronométré par un « timer » inséré dans chacune des pages de l'enquête en ligne. En effet, un temps de réponse très faible trahit le fait que le répondant n'a pas pris le temps de lire les questions avant d'y répondre et qu'il a coché les modalités de réponse aléatoirement (Parguel, 2009).

- Les réponses incohérentes, c'est-à-dire toutes les réponses données sans prêter attention au contenu de l'item (Graham, 1993, p. 38). Il peut s'agir d'une réponse identique à deux items inversés ou de la sélection systématique de la même réponse sur une page entière de questionnaire.

Le taux de réponses aléatoires n'était pas significativement différent entre les deux conditions (PR_{ESG} = 23,5 % de réponses aléatoires vs. PR_{Ds} = 20 % de réponses aléatoires ; p = 0,47).

Suite à cette phase d'épuration, l'échantillon se constituait de 313 répondants, soit 153 pour la *condition ESG* et 160 pour la *condition DS*. Les deux groupes expérimentaux étaient homogènes en termes de sexe (PR_{ESG} = 53,6 % femmes vs. PR_{DS} = 50 % femmes ; p = 0,57) et tous les répondants avaient entre 30 et 40 ans, tranche d'âge susceptible de garantir l'implication personnelle des participants dans une relation avec une banque (M_{ESG} = 35,6 ans vs. M_{DS} = 35,2 ans ; p = 0,32).

4. Format du questionnaire

Deux versions de l'enquête en ligne ont été développées. Pour nous prémunir contre un éventuel effet de contamination entre les questions, l'évaluation globale a été placée en premier, avant la qualité perçue de chacun des attributs du service bancaire (Bickart, 1993). Dans cette perspective, la première page de l'étude concernait la mesure de la satisfaction globale, tandis que les pages suivantes portaient sur la perception de la qualité du service bancaire. La mesure des intentions comportementales se situait à la fin de l'étude, juste avant les renseignements sociodémographiques généraux sur le répondant. La seule différence entre les deux versions du questionnaire était l'outil de mesure utilisé pour compléter les items de satisfaction. Dans la *condition ESG*, une brève explication d'utilisation précédait la présentation de la première grille. Les deux versions du questionnaire sont présentées en annexe 3.

5. Echelles de mesure utilisées

Toutes les échelles employées pour mesurer les construits mobilisés dans le modèle théorique ont été validées empiriquement dans la littérature. Pour les besoins de la collecte, elles sont été traduites de l'anglais vers le français.

5.1. Satisfaction

Pour ne pas alourdir notre protocole, nous avons opté pour une mesure directe de la satisfaction, considérant, comme les tenants du paradigme « *performance-based* » (Cronin Jr et Taylor, 1992), que les réponses exprimées par les sujets prennent directement en compte leurs attentes. Pour capturer ce construit, nous avons utilisé une échelle multi-items composée de 3 paires d'adjectifs antonymes très acceptées pour mesurer la satisfaction (Anderson et Srinivasan, 2003 ; Evanschitzky et Wunderlich, 2006 ; Lam & al., 2004 ; Mittal, Ross et Baldasare, 1998). Les répondants indiquaient leur niveau de satisfaction globale à l'égard de leur banque, soit sur des *Evaluative Space Grids* en 25 points (*condition ESG*), soit sur des différentiels sémantiques en 9 points (*condition DS*). Le choix d'un différentiel sémantique en 9 points a été guidé par la perspective de création d'un index en 9 points à partir des évaluations effectuées sur la grille de Larsen (voir développement de la section 3. sur le codage de la grille). Les items de l'échelle correspondaient aux trois paires d'antonymes suivantes : « opinions extrêmement négatives » / « opinions extrêmement positives » ; « extrême insatisfaction » / « extrême satisfaction » et « opinions extrêmement défavorables » / « opinions extrêmement favorables ».

5.2. Qualité perçue

La qualité perçue correspond à l'opinion à l'égard des différents attributs de la performance (Oliver, 1997). Nous avons utilisé l'échelle de mesure de la qualité des banques de Karatepe et ses collègues (2005) pour évaluer la performance perçue par les répondants de leur établissement financier. Il s'agit d'une adaptation de l'approche SERVQUAL à l'évaluation du service bancaire (Parasuraman, Zeithaml et Berry, 1988). Cette échelle est composée de 20 items qui se regroupent en quatre dimensions : l'environnement du service (4 items), la qualité des interactions (7 items), l'empathie (5 items) et la fiabilité (4 items). Les items étaient mesurés sur des échelles de Likert en 7 points (allant de 1= « pas du tout d'accord » à 7= « tout à fait d'accord ») (Tableau 19).

5.3. Intentions comportementales

Pour examiner la validité prédictive de l'*Evaluative Space Grid* comme outil de mesure de la satisfaction globale, nous avons testé sa capacité à prédire deux variables d'intentions fondamentales en marketing : l'intention de fidélité et l'intention de bouche-à-oreille. L'intention des répondants d'être fidèles envers leur banque a été mesurée à l'aide des trois

Construits et items	Moy.	Ecart type	Moy.	Ecart type
Qualité perçue				
Environnement du service	*Condition DS*		*Condition ESG*	
L'extérieur de votre agence bancaire est attrayant	4,21	1,58	4,33	1,39
L'intérieur de votre agence bancaire est attrayant	4,43	1,47	4,54	1,34
Les employés de cette banque ont une apparence propre et soignée	5,43	1,32	5,43	1,09
L'intérieur de votre agence bancaire est spacieux	4,61	1,55	4,58	1,46
Qualité des interactions	*Condition DS*		*Condition ESG*	
Les employés de cette banque ont les connaissances nécessaires pour répondre aux problèmes	4,88	1,33	4,86	1,41
Les employés de cette banque sont polis avec les clients	5,62	1,18	5,47	1,20
Les employés de cette banque sont expérimentés	5,04	1,20	4,92	1,29
Les employés de cette banque inspirent la confiance aux clients	4,97	1,31	4,88	1,33
Les employés de cette banque sont à l'écoute des clients	5,16	1,28	5,03	1,34
Les employés de cette banque servent les clients de manière satisfaisante	5,17	1,26	5,05	1,26
Les employés de cette banque sont chaleureux avec leurs clients	4,95	1,34	5,05	1,31
Empathie	*Condition DS*		*Condition ESG*	
Cette banque ne laisse pas ses clients faire la queue pendant un temps interminable	4,73	1,57	4,67	1,45
Les employés de cette banque réalisent les opérations en temps voulu	5,03	1,44	5,03	1,31
Les employés de cette banque aident systématiquement les clients	4,94	1,44	4,87	1,33
Les employés de cette banque fournissent une attention personnalisée aux clients	4,81	1,41	4,86	1,34
Les employés de cette banque sont disposés à résoudre les problèmes des clients	4,94	1,40	4,93	1,30
Fiabilité	*Condition DS*		*Condition ESG*	
Les employés de cette banque ne font pas d'erreur	4,51	1,36	4,35	1,50
Les employés de cette banque respectent la confidentialité des transactions	5,53	1,22	5,42	1,17
Les employés de cette banque fournissent des informations précises aux clients	5,02	1,27	4,92	1,37
Cette banque informe les clients de leurs opérations financières avec précision	5,01	1,32	4,95	1,32
Satisfaction globale	*Condition DS*		*Condition ESG*	
Opinions extrêmement négatives - Opinions extrêmement positives	6,10	1,85	5,29	1,55
Extrême insatisfaction - Extrême satisfaction	6,08	1,74	5,41	1,59
Opinions extrêmement défavorables - Opinions extrêmement favorables	6,06	1,90	5,51	1,52
Intentions de fidélité	*Condition DS*		*Condition ESG*	
J'ai l'intention d'utiliser les services de cette banque dans les années à venir	5,23	1,36	4,93	1,53
Je vais utiliser les services de cette banque dans les années à venir	5,03	1,44	4,77	1,61
Je souhaite utiliser les services de cette banque dans les années à venir	5,04	1,44	4,76	1,64
Intentions de bouche-à-oreille	*Condition DS*		*Condition ESG*	
J'ai l'intention de dire des choses positives sur cette banque dans les années à venir	4,43	1,54	4,35	1,51
Je vais encourager mes amis et mes proches à faire appel à cette bqe	4,16	1,58	3,76	1,61
Je vais recommander cette banque aux personnes qui me demandent conseil	4,48	1,63	4,29	1,68

Tableau 19. Moyenne et écart-type pour chaque item du questionnaire par condition

items proposés par Seiders et ses collègues (2005) pour capturer la fidélité comportementale : « J'ai l'intention d'utiliser les services de cette banque dans les années à venir », « Je vais utiliser les services de cette banque dans les années à venir » et « Je souhaite utiliser les services de cette banque dans les années à venir ». Ces items ont été évalués sur des échelles de Likert en 7 points (allant de 1= « très peu probable » à 7 = « très probable »). Les intentions de bouche-à-oreille positif ont été mesurées avec les trois items proposés par Zeithaml et ses collègues (1996) : « J'ai l'intention de dire des choses positives aux gens sur cette banque dans les années à venir », « Je vais encourager mes amis et mes proches à faire appel à cette banque », « Je vais recommander cette banque aux personnes qui me demandent conseil ». Ces items ont été évalués sur des échelles de Likert en 7 points (allant de 1= « pas du tout d'accord » à 7 = « tout à fait d'accord »).

Le tableau 19 présente les moyennes et les écart-types pour chaque item. Il n'y a pas de différence significative entre les conditions expérimentales.

Section 3. Analyses et résultats

1. Procédure de codage préalable

L'objectif d'une mesure est de classer les individus (Thornton, 2011). Dans le cas présent, il s'agit de les ordonner suivant deux critères : leur niveau de satisfaction et leur niveau d'ambivalence. Or, l'*Evaluative Space Grid* n'ayant jamais été employée dans le cadre de la mesure de la satisfaction, nous avons dû définir un codage. Le premier objectif de ce codage était de fournir un index permettant la comparaison des notes entre les deux conditions. Nous avons volontairement mesuré la satisfaction globale sur des échelles en 9 point dans la *condition DS* car la grille de Larsen et ses collègues (2009), matrice 5 x 5, correspond à la réunion de deux continuums gradués de 0 à 4. En effet, une rotation à 90 degrés de l'axe vertical de la négativité permet de transformer une *Evaluative Space Grid* en un différentiel sémantique en 9 points, comme le montre la figure 25 (page 104).

Le deuxième objectif de cette procédure de codage était de classer les individus suivant leur niveau de satisfaction globale. Dans cette perspective, nous ne pouvions pas utiliser la formule proposée par Thompson et ses collègues (1995) puisqu'elle a uniquement pour objectif de capturer l'ambivalence. Pour rappel, cette équation « valorise » les situations de co-activation de réactions positives et négatives et « pénalise » les évaluations polarisées

(cf. Chapitre 1, Section 2). Cette formule répond donc à une logique différente de la nôtre dans la mesure où nous souhaitons créer un codage permettant de distinguer les quatre réponses évaluatives décrites précédemment : satisfaction, insatisfaction, indifférence et ambivalence. Il s'agit donc pour nous de valoriser à la fois les situations de co-activation de réactions positives et négatives et les évaluations polarisées.

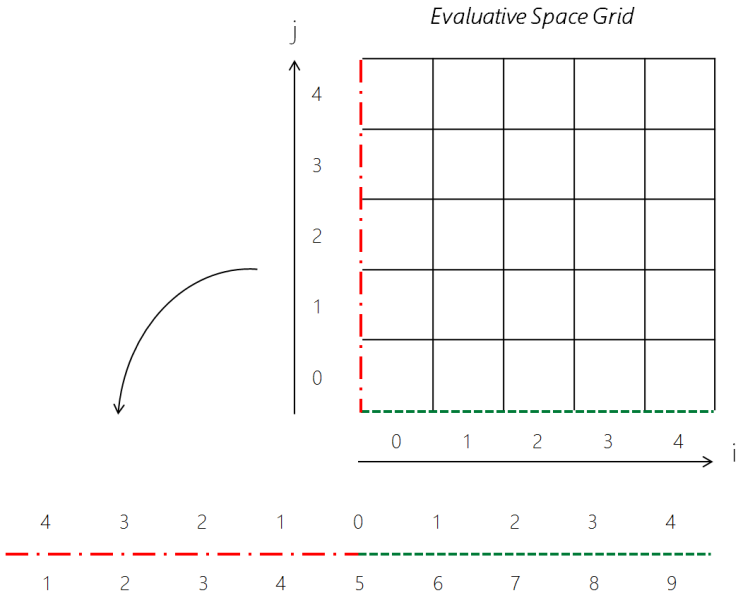


Figure 26. Transformation d'une grille de Larsen en différentiel sémantique en 9 points

Suivant la démarche adoptée par Breckler (1994) ou par Thompson et ses collègues (1995) pour discuter la pertinence des différents index de mesure de l'ambivalence, nous avons proposé une formule permettant de calculer un score de satisfaction globale (S) à partir de la note attribuée sur l'axe horizontal de mesure de la satisfaction, « i » et sur l'axe vertical de mesure de l'insatisfaction, « j ». Pour des raisons de simplicité, nous avons opté pour un modèle bilinéaire du type[9] :

$$S(i;j)=ai+bj+c$$

a, b et c étant des paramètres constants et indépendants de i et j

[9] *Nous tenons à remercier très sincèrement Dominique Mahut, ingénieur d'études statistiques au CNRS, pour sa contribution décisive au développement de cette formule.*

Comme l'avaient faits Breckler (1994) ou Thompson et collègues (1995), nous avons ensuite défini les propriétés désirables de l'index de satisfaction global. Dans le but, d'estimer les paramètres du modèle bilinéaire, ces propriétés ont été traduites en contraintes mathématiques. Cette démarche nous permis de dégager les six contraintes suivantes :

- Le score de satisfaction globale, S(i,j), doit diminuer à mesure que la perception d'insatisfaction augmente le long de l'axe vertical de mesure de l'insatisfaction (j). Ainsi, l'insatisfaction maximale est atteinte pour la cellule située tout en haut à gauche de la grille, correspondant à une évaluation « extrêmement insatisfaite » et « pas du tout satisfaite ». Cette cellule de la grille est codée 1.

  ➔ *Contrainte 1 : S(i,j) diminue à « i » constant si « j » augmente*

  ➔ *Contrainte 2 : S(1;5) = 1 = minimum de S(i;j)*

- Le score de satisfaction globale, S(i,j), doit augmenter à mesure que la perception de satisfaction augmente le long de l'axe horizontal de mesure de la satisfaction (i). Ainsi, la satisfaction maximale est atteinte pour la cellule située tout en bas à droite de la grille, correspondant à une évaluation « extrêmement satisfaite » et « pas du tout insatisfaite ». Cette cellule de la grille est codée 9.

  ➔ *Contrainte 3 : S(i,j) augmente à « j » constant si « i » augmente*

  ➔ *Contrainte 4 : S(5 ;1) = 9 = maximum de S(i;j)*

- Chaque point de l'index de satisfaction globale, S(i,j), doit se traduire par des scores différents de manière à ce que deux répondants choisissant des cellules différentes se voient attribués des notes différentes :

  ➔ Contrainte 5 : S(i,j) injective, soit S(i,j) = S(i', j') implique i = i' et j = j'

- Quand les perceptions de satisfaction et d'insatisfaction sont concomitantes, l'ambivalence doit augmenter à mesure que la satisfaction et l'insatisfaction augmentent. De manière réciproque, l'indifférence doit augmenter à mesure que la satisfaction et l'insatisfaction diminuent. Ainsi, l'ambivalence maximale est atteinte pour la cellule située tout en haut à droite de la grille, correspond à une évaluation « extrêmement satisfaite » et « extrêmement insatisfaite ». Cette cellule de la grille est codée 7. De même, l'indifférence maximale est atteinte pour la cellule située tout en bas à gauche de la grille, correspondant à une évaluation «

pas du tout satisfaite » et « pas du tout insatisfaite ». Cette cellule de la grille est codée 3.

➔ *Contrainte 6 : g(i) = S(i,j) croissant*

La traduction de ces contraintes s'est répercutée de la manière suivante au niveau des paramètres du modèle :

Contrainte à respecter	Conséquences sur l'estimation des paramètres
Contrainte 1	b<0
Contrainte 3	a>0
Contraintes 2 et 4	a = b+2 ; c = -1-6b
Contrainte 5	Le modèle retenu ne permet pas de respecter la contrainte 5
Contrainte 6	b>-1

Tableau 20. Impact des contraintes définies sur l'estimation des paramètres du modèle

La synthèse de ces contraintes aboutie à la formule suivante :

$$S(i,j) = (b+2)i + bj - 1 - 6b$$

Avec $-1<b<0$

Nous avons ensuite testé différentes valeurs de « b » dans le but d'optimiser les contraintes. Pour chaque valeur testée, un indice de pertinence, traduisant le degré d'adéquation aux contraintes, a été calculé. Nous avons finalement opté pour b = -0,5, valeur permettant d'atteindre un indice de pertinence de 0,9756. Cette valeur offre également l'avantage de faciliter la comparaison entre l'*Evaluative Space Grid* et le différentiel sémantique, puisque le point central de la grille, correspondant à une évaluation « modérément satisfaite » et « modérément insatisfaite » se voit attribuer le score de 5, tout comme le point central du différentiel sémantique.

L'application de cette formule nous a permis de proposer l'index de satisfaction globale présenté en figure 27. C'est cet index qui a été utilisé comme codage lors des analyses réalisées dans le cadre de cette étude.

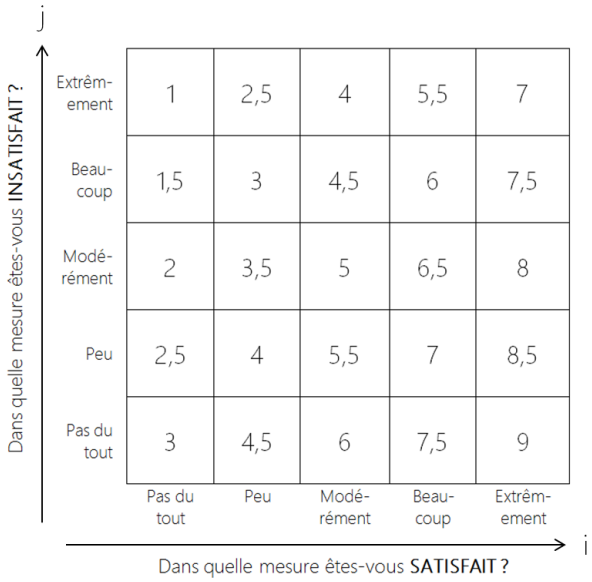


Figure 27. Codage de l'*Evaluative Space Grid* utilisé pour l'analyse

Dans la formule retenue, le paramètre a (=1,5), associé à la satisfaction est supérieur au paramètre b (=-0,5), associé au score d'insatisfaction. Cette asymétrie de poids entre la satisfaction et l'insatisfaction était nécessaire pour développer un score ascendant sur la diagonale de la grille (contrainte 6). Si ce codage peut sembler quelque peu arbitraire, il est cohérent avec les recherches portant sur la distribution asymétrique négative de la satisfaction (Hurley et Estelami, 1998 ; Peterson et Wilson, 1992). Comme nous l'avons déjà évoqué, ces recherches montrent que la distribution de la satisfaction ne suit pas une loi normale et se caractérise par une distribution asymétrique négative. Cela se traduit par des valeurs basses sur la partie gauche de la queue de la distribution, c'est-à-dire par un biais de positivité (Estelami et De Maeyer, 1997 ; Hurley et Estelami, 1998 ; Parasuraman, Zeithaml et Berry, 1994 ; Peterson et Wilson, 1992 ; Westbrook, 1980). A l'appui de ces travaux, le codage que nous proposons accorde davantage de place à la satisfaction qu'à l'insatisfaction en condensant les notes attribuées à l'insatisfaction et en étendant les notes attribuées à la satisfaction.

2. Analyse des données

Les deux modèles de mesure alternatifs (*Evaluative Space Grids versus* différentiels sémantiques) ont été testés en recourant à la modélisation par analyses structurelles. Le choix des méthodes d'équations structurelles se justifie de plusieurs manières. En effet,

selon Roussel et ses collègues (2002, p.14) « ces modèles [...] permettent d'introduire simultanément plusieurs variables à expliquer dans une même analyse ». C'est précisément ce que nous cherchons à faire dans la mesure où nous souhaitons comparer la performance de deux outils de mesure pour prédire les variations de deux variables dépendantes : les intentions de fidélité et les intentions de bouche-à-oreille. Toujours selon Roussel et ses collègues (2002, p.15), les modèles d'équations structurelles « permettent de tester non seulement les effets de variables prises isolément sur d'autres variables mais aussi les effets conjoints de plusieurs variables sur une ou plusieurs autres ». C'est le cas dans notre modèle théorique qui postule un effet associé de la qualité perçue et de la satisfaction globale sur les intentions comportementales. Roussel et ses collègues (2002, p.15) ajoutent que « ces modèles offrent la possibilité d'examiner des variables intermédiaires, qui ont le statut de variables dépendantes pour plusieurs relations dans un modèle, et celui de variables indépendantes pour d'autres ». Dans le modèle illustré par la figure 25 (page 104) la variable « satisfaction globale » a ce double statut puisqu'il s'agit d'une variable dépendante du point de vue de la « qualité perçue » et d'une variable explicative des « intentions de fidélité » et des « intentions de bouche-à-oreille ».

Pour tester notre modèle, nous avons suivi une démarche en trois temps. Il s'agit d'abord de mener des analyses exploratoires sur les données relatives aux différents construits du modèle théorique (2.1.). Cette première étape est suivie par une phase d'analyses confirmatoires visant à nous assurer de la validité de trait des construits, ainsi qu'à comparer la structure factorielle de nos deux modèles de mesure (2.2.). Enfin, nous avons procédé au test effectif du modèle structurel (2.3.).

2.1. Analyses factorielles exploratoires

L'analyse exploratoire des données a pour visée de vérifier la fiabilité et la dimensionnalité des construits composant le modèle de recherche. Nous avons effectué des analyses exploratoires en axes principaux, variable par variable, en nous assurant de la dimensionnalité des construits utilisés et de la qualité de représentation des items. Pour interpréter ces résultats, nous nous sommes appuyés sur les indices suivants :

- L'indice de Kaiser-Meyer-Oklin (KMO) : il permet de juger du degré d'inter-corrélation entre les items d'un même construit. Lorsqu'il est compris entre 0,7 et 1, on reconnaît le caractère factorisable des données (Evrard, Pras et Roux, 2003).

- Les contributions factorielles et les communalités évaluent la qualité des items. Ces critères permettent d'épurer les échelles en supprimant ceux qui ne contribuent pas à expliquer le construit.

    - Les contributions factorielles permettent d'ordonner les axes factoriels selon leur capacité à restituer l'information. Selon Hair et ses collègues (2006), la contribution factorielle doit au moins être égale à 0,4 pour que l'item soit retenu.

    - Les communalités correspondent au pourcentage de variance expliquée dans les dimensions extraites. Elles doivent également être supérieures à 0,4 pour que l'item soir retenu.

- L'alpha de Cronbach est le test de fiabilité des échelles de mesure le plus commun. Il s'agit d'un indice statistique variant entre 0 et 1 qui permet d'évaluer l'homogénéité et la cohérence interne d'un instrument. On considère généralement que la fiabilité de l'instrument est satisfaisante lorsque la valeur du coefficient est au moins égale à 0,7 (Roussel & al., 2002).

Compte tenu de notre objectif de comparaison entre deux modèles de mesure, nous avons également veillé à ce que nos deux ensembles de données restent comparables. En conséquence, nous avons fait des choix permettant de conserver les mêmes items dans les deux ensembles de données.

Nous avons commencé par examiner la dimensionnalité de l'échelle de la qualité des banques de Karatape et ses collègues (2005). Pour les deux ensembles de données, la solution initiale comporte seulement deux axes, le premier axe étant alimenté par la grande majorité des items. Le second axe est alimenté par deux items (*condition ESG*) ou trois items (*condition DS*) de la dimension environnement du service qui, de plus, présentent des contributions factorielles significatives sur le premier axe. Dans la *condition DS*, le premier axe permet de restituer 56,2% de l'information et le second 7,9%. Dans la *condition ESG*, le premier axe permet de restituer 56,5% de l'information et le second 8,4%. A l'exception de l'item « L'extérieur de votre agence bancaire est attrayant » qui, dans la *condition ESG*, ne contribue qu'au deuxième axe, tous les items ont une contribution factorielle supérieure à 0,4 sur le premier axe. Cette première analyse nous impose d'exclure la dimension environnement du service de l'échelle de la qualité des banques et de ne retenir qu'un seul facteur. Cette solution n'est pas satisfaisante par rapport à nos objectifs. Dans une

perspective d'étude de l'ambivalence, il nous apparait important de conserver les quatre dimensions proposées par les auteurs. En effet, certaines pourraient être évaluées positivement et d'autres négativement dans le cas d'une évaluation ambivalente. Nous avons donc décidé d'analyser séparément les caractéristiques des quatre dimensions du construit de qualité perçue avec la perspective de recourir à une variable de second ordre dans la suite des analyses.

Cette analyse nous a conduit à supprimer l'item « les employés de cette banque ont une apparence propre et soignée » de la dimension « environnement du service », ainsi que l'item « cette banque ne laisse pas ses clients faire la queue pendant un temps interminable » de la dimension « empathie » et l'item « les employés de cette banque respectent la confidentialité des transactions » de la dimension fiabilité. En effet, pour ces trois indicateurs, les communalités étaient à chaque fois inférieures à 0,4 dans les deux ensembles de données. Tous les items de la dimension « qualité des interactions » ont été conservés car ils présentaient des caractéristiques satisfaisantes dans les deux conditions expérimentales. Après cette première phase d'épuration, l'item « l'intérieur de votre agence bancaire est spacieux » contribuait faiblement à la dimension « environnement du service » dans les deux conditions. Nous avons toutefois décidé de le conserver dans le but de garder un minimum de trois items par dimension (Little, Lindenberger et Nesselroade, 1999 ; Velicer et Fava, 1998) et parce que sa suppression détériorait l'alpha de Cronbach.

Le tableau 20 présente les résultats après épuration de l'échelle.

Qualité des banques				
Construits et items	Com.	C. F.	Com.	C.F.
Environnement du service	*Condition DS*		*Condition ESG*	
L'extérieur de votre agence bancaire est attrayant	0,64	0,80	0,70	0,84
L'intérieur de votre agence bancaire est attrayant	0,92	0,96	0,89	0,94
L'intérieur de votre agence bancaire est spacieux	0,36	0,60	0,24	0,50
Valeur propre	1,92		1,83	
% de la variance expliquée	64,03		61,09	
KMO	0,65		0,62	
Alpha de Cronbach	0,82		0,79	
Qualité des interactions	*Condition DS*		*Condition ESG*	
Les employés de cette banque ont les connaissances nécessaires pour répondre aux problèmes	0,68	0,82	0,71	0,84
Les employés de cette banque sont polis avec les clients	0,57	0,77	0,58	0,76
Les employés de cette banque sont expérimentés	0,68	0,83	0,83	0,91
Les employés de cette banque inspirent la confiance aux clients	0,84	0,92	0,82	0,91
Les employés de cette banque sont à l'écoute des clients	0,77	0,88	0,83	0,91
Les employés de cette banque servent les clients de manière satisfaisante	0,74	0,86	0,82	0,91
Les employés de cette banque sont chaleureux avec leurs clients	0,70	0,84	0,68	0,82
Valeur propre	5,01		5,28	
% de la variance expliquée	71,51		75,44	
KMO	0,92		0,93	
Alpha de Cronbach	0,95		0,95	
Empathie	*Condition DS*		*Condition ESG*	
Les employés de cette banque réalisent les opérations en temps voulu	0,71	0,84	0,64	0,80
Les employés de cette banque aident systématiquement les clients	0,85	0,92	0,73	0,86
Les employés de cette banque fournissent une attention personnalisée aux clients	0,78	0,88	0,71	0,84
Les employés de cette banque sont disposés à résoudre les problèmes des clients	0,81	0,90	0,85	0,92
Valeur propre	3,15		2,93	
% de la variance expliquée	78,82		73,24	
KMO	0,84		0,84	
Alpha de Cronbach	0,94		0,92	
Fiabilité	*Condition DS*		*Condition ESG*	
Les employés de cette banque ne font pas d'erreur	0,61	0,78	0,43	0,65
Les employés de cette banque fournissent des informations précises aux clients	0,91	0,90	0,91	0,95
Cette banque informe les clients de leurs opérations financières avec précision	0,669	0,95	0,66	0,81
Valeur propre	2,32		1,99	
% de la variance expliquée	77,47		66,27	
KMO	0,72		0,67	
Alpha de Cronbach	0,83		0,91	

Com. : communalités ; C.F. : contributions factorielles

Tableau 21. Résultats suite à l'épuration de l'échelle de la qualité des banques

En ce qui concerne les autres construits du modèle théorique, les résultats se sont avérés satisfaisants. Tous les construits sont unidimensionnels et présentent des indices de fiabilité supérieurs 0,7. Le détail de ces analyses figure dans le tableau 21.

Satisfaction globale	Com.	C.F.	Com.	C.F.
	Condition DS		*Condition ESG*	
Opinions extrêmement négatives - Opinions extrêmement positives	0,96	0,98	0,72	0,85
Extrême insatisfaction - Extrême satisfaction	0,85	0,92	0,75	0,87
Opinions extrêmement défavorables - Opinions extrêmement favorables	0,42	0,65	0,82	0,91
Valeur propre	2,22		2,29	
% de la variance expliquée	74,08		76,38	
KMO	0,66		0,75	
Alpha de Cronbach	0,88		0,9	
Intentions de fidélité	**Com.**	**CF.**	**Com.**	**CF.**
	Condition DS		*Condition ESG*	
J'ai l'intention d'utiliser les services de cette banque dans les années à venir	0,89	0,94	0,8	0,9
Je vais utiliser les services de cette banque dans les années à venir	0,93	0,97	0,91	0,95
Je souhaite utiliser les services de cette banque dans les années à venir	0,93	0,96	0,94	0,97
Valeur propre	2,75		2,65	
% de la variance expliquée	91,72		88,48	
KMO	0,76		0,76	
Alpha de Cronbach	0,96		0,96	
Intentions de bouche-à-oreille	**Com.**	**CF.**	**Com.**	**CF.**
	Condition DS		*Condition ESG*	
J'ai l'intention de dire des choses positives sur cette banque dans les années à venir	0,77	0,88	0,7	0,84
Je vais encourager mes amis et mes proches à faire appel à cette banque	0,83	0,91	0,76	0,87
Je vais recommander cette banque aux personnes qui me demandent conseil	0,92	0,96	0,76	0,87
Valeur propre	2,52		2,22	
% de la variance expliquée	84,14		74,06	
KMO	0,68		0,75	
Alpha de Cronbach	0,91		0,9	

Com. : communalités ; C.F. : contributions factorielles

Tableau 22. Résultats des analyses exploratoires pour la satisfaction globale, les intentions de fidélité et les intentions de bouche-à-oreille

2.2. Analyses factorielles confirmatoires

Dans le prolongement des analyses factorielles exploratoires, les analyses factorielles confirmatoires s'intéressent à la fois à la fiabilité et à la validité des mesures. Nous avons utilisé le logiciel LISREL pour effectuer ces analyses séparément pour chaque ensemble de données (*condition ESG* et *condition DS*).

Fiabilité et validité convergente

Pour apprécier la cohérence interne des construits, nous avons calculé le rhô de Jöreskog. Ce coefficient s'interprète de la même manière que l'alpha de Cronbach : la fiabilité du construit est bonne si l'indice est supérieur au seuil critique de 0,7. Toutefois, le rhô de Jöreskog diffère de l'alpha de Cronbach dans la mesure où il intègre dans son calcul les termes d'erreurs et s'avère insensible au nombre d'items utilisés. Les résultats, présentés dans le tableau 22, indiquent une bonne cohérence interne entre les items dans les deux ensembles de données, dans la mesure où les rhôs des échelles sont tous supérieurs à 0,70.

Construits	Rhô de Jöreskog	Validité conv.	Rhô de Jöreskog	Validité conv.
	Condition DS		*Condition ESG*	
Environnement du service	0,97	0,91	0,96	0,90
Qualité des interactions	0,99	0,92	0,99	0,92
Empathie	0,98	0,93	0,98	0,92
Fiabilité	0,98	0,93	0,97	0,91
Satisfaction globale	0,97	0,92	0,81	0,81
Intention de fidélité	0,98	0,94	0,84	0,84
Intention de bouche-à-oreille	0,84	0,84	0,97	0,92

Validité conv. : variance moyenne extraite

Tableau 23. Indicateurs de fiabilité et de validité convergente des construits

L'appréciation de la validité convergente consiste à vérifier que les items ou les indicateurs d'un même construit sont corrélés entre eux. Elle est confirmée lorsque les tests de Student associés à chaque contribution factorielle sont significatifs (Fornell et Larcker, 1981). Un autre moyen de s'assurer de la validité convergente d'une mesure est de vérifier que la moyenne des variances extraites partagées entre le construit et ses mesures est supérieure à 0,50. Le rhô de validité convergente est l'indicateur employé à cet effet. Les résultats

indiquent que dans les deux conditions, d'une part tous les items ont des contributions factorielles significatives supérieures à 0,70 et d'autre part, chaque construit partage plus de variance avec ses mesures qu'avec l'erreur. En effet, tous les rhôs de validité convergente sont supérieurs à 0,5. Les résultats attestent de la validité convergente des échelles de mesure (tableau 22).

Validité discriminante

Pour apprécier la validité discriminante de nos deux modèles de mesure, nous avons comparé plusieurs modèles nichés comportant de manières séparée ou agrégée les mesures des différents construits. L'objectif de cette démarche est de préciser, sur la base d'une série de tests de différence de Khi-deux, les qualités d'ajustement de nos modèles de mesure et en conséquence, la validité des construits utilisés. Nous avons cherché à nous assurer que les différents construits du modèle sont bien distincts les uns des autres. Compte tenu de la faible taille de l'échantillon effectif, nous avons opté pour le test de la validité discriminante et du modèle étape par étape (Hair & al., 2006 ; Roussel et Wacheux, 2005). Il a donc fallu opérer des choix. Pour nous aider à faire un choix quant aux variables à combiner et aux modèles à tester, nous nous sommes appuyés, d'une part, sur les corrélations existant entre les variables en question, et d'autre part, sur leur proximité conceptuelle. Nous avons ainsi testé la validité discriminante entre les construits de qualité perçue et de satisfaction d'une part, et entre l'intention de bouche-à-oreille et l'intention de fidélité d'autre part.

Le premier volet concerné par ces analyses se rapporte la qualité perçue et à la satisfaction globale. En opérant des choix quant aux construits à combiner, nous avons mis en concurrence trois modèles. Le détail de la composition de ces modèles et les indices d'ajustements qui leur sont associés figurent dans le tableau 23. Dans les deux ensembles de données, c'est le premier modèle qui présente les meilleurs indices d'ajustement. Nous retenons donc la solution distinguant une variable du second ordre de la qualité perçue combinant les quatre dimensions proposées par Karatepe et ses collègues (Karatepe, Yavas et Babakus) et la satisfaction globale. Ce modèle est significativement meilleur que les deux autres puisque, la différence des Khi-deux rapportée à la différence des degrés de liberté est significative à p> 0,001.

Modèles	χ_2	ddl	$\Delta\chi_2$	Δddl	AIC	RMSEA	CFI	GFI
			Condition DS					
Modèle 1	345,271**	157	-	-	441,271	0,093	0,941	0.822
Modèle 2	424,450	160	79,18**	3	524,450	0,102	0,924	0.789
Modèle 3	1161,360	170	736,91**	10	1241,360	0,192	0,727	0.578
			Condition ESG					
Modèle 1	352,815**	157	-	-	458,809	0,090	0,940	0,812
Modèle 2	464,218	160	111,403**	3	564,208	0,112	0,908	0,766
Modèle 3	916,792	170	452,574**	10	996,792	0,170	0,775	0,624

- Modèle 1 : Distinction entre une variable du second ordre de qualité perçue combinant les quatre dimensions proposées par Karatepe et ses collègues (Karatepe, Yavas et Babakus) et la satisfaction globale.

- Modèle 2 : Distinction entre les quatre composantes de la qualité perçue (environnement du service, qualité des interactions, empathie et fiabilité) et la satisfaction globale.

- Modèle 3 : Les construits de qualité perçue et de satisfaction globale sont confondus de manière à ne former qu'un seul construit.

Tableau 24. Indices d'ajustement pour les modèles alternatifs testés entre la satisfaction globale et qualité perçue

Le deuxième volet concerné par ces analyses se rapporte aux intentions de fidélité et de bouche-à-oreille. Pour vérifier que ces deux variables dépendantes correspondaient bien à deux construits distincts, nous avons mis en concurrence deux modèles. Le détail de la composition de ces modèles et les indices d'ajustement qui leur sont associés figurent dans le tableau 24. Dans les deux ensembles de données, c'est le premier modèle qui présente les meilleurs indices d'ajustement. Nous retenons donc la solution envisageant les intentions de fidélité comme un construit distinct des intentions de bouche-à-oreille (modèle 1). Ce modèle est significativement meilleur que le second puisque la différence des Khi-deux rapportée à la différence des degrés de liberté est significative à p> 0,001.

Modèles	χ_2	ddl	$\Delta\chi_2$	Δddl	AIC	RMSEA	CFI	GFI
			Condition DS					
Modèle 1	11,452**	7	-	-	41,452	0,076	0,995	0.977
Modèle 2	172,839	9	161,387**	2	196,839	0,338	0,852	0.734
			Condition ESG					
Modèle 1	12,040**	7	-	-	40,040	0,069	0,996	0,974
Modèle 2	110,864	9	98,824**	2	134,864	0,273	0,920	0,804

- Modèle 1 : Les construits d'intention de fidélité et d'intention de bouche-à-oreille sont distincts.

- Modèle 2 : Les construits d'intention de fidélité et d'intention de bouche-à-oreille sont confondus.

Tableau 25. Indices d'ajustement pour les modèles alternatifs testés entre les intentions de fidélité et les intentions de bouche-à-oreille

2.3. Test du modèle structurel

Après avoir apprécié la fiabilité et les validités convergente et discriminante des construits de nos deux modèles de mesure, l'ultime étape de l'analyse consiste à tester l'ajustement des modèles de mesure dans leur ensemble au modèle théorique. Cette procédure permet également de se pencher sur le pouvoir prédictif des deux modèles alternatifs de mesure (ESG et DS).

Pour apprécier les qualités d'ajustement de nos modèles et les comparer entre eux, nous retenons quatre indices d'ajustement d'un modèle de mesure à un modèle théorique (Roussel et al., 2002) :

- Un indice absolu, le RMSEA (« Root Mean Square Error of Approximation ») : l'ajustement est jugé acceptable dès lors que cet indice est inférieur à 0,08, mais il est préférable de retenir les modèles dont le RMSEA est inférieur à 0,05 ;
- Deux indices incrémentaux :
    - Le CFI (« Comparative Fit Index ») : ce coefficient est jugé satisfaisant dès lors qu'il atteint ou dépasse 0,9 ;
    - Le GFI « goodness-of-fit » : il doit être supérieur à 0,7 pour que l'ajustement au modèle soit jugé satisfaisant.
- Un indice de parcimonie, l'AIC, qui permet la comparaison entre deux modèles de mesure. Il n'a de sens que lorsque plusieurs modèles sont estimés. Plus il est faible, meilleur est l'ajustement au modèle.

Notre modèle théorique envisage la satisfaction globale comme un médiateur de la relation entre la qualité perçue et les intentions comportementales étudiées. L'hypothèse stipule que cette médiation est partielle puisque des liens directs entre la qualité perçue et les variables dépendantes du modèle sont envisagés. De plus, une relation positive entre les intentions de fidélité et de bouche-à-oreille est également envisagée (Anderson, Fornell et Lehmann, 1994 ; Cronin, Brady et Hult, 2000 ; Szymanski et Henard, 2001).

Les résultats montrent que l'ajustement au modèle théorique est acceptable pour les deux modèles de mesure. Les indicateurs de la qualité d'ajustement (CFI et GFI) sont au-delà des seuils d'acceptabilité et les AIC sont globalement comparables entre les modèles. Seul le

RMSEA ne satisfait pas le seuil d'acceptabilité de 0,8 pour les deux modèles. Cela peut s'expliquer par la petite taille de nos échantillons (Hu et Bentler, 1998 ; Jolibert et Jourdan, 2005). Les valeurs des indicateurs sont présentées dans le tableau 25.

Indicateurs d'ajustement	Seuils	*Condition DS*	*Condition ESG*
RMSEA	< 0,08 et si possible < 0,05	0,090	0,092
CFI	> 0,9	0,913	0,915
GFI	> 0,7	0,761	0,756
AIC	Le plus faible possible	786,166	784,597

Tableau 26. Indicateurs d'ajustement associés aux deux modèles de mesure

Ces indicateurs étant sensiblement identiques dans les deux conditions expérimentales, il nous est difficile, à ce stade de juger de la supériorité de l'un par rapport à l'autre. Au-delà de l'ajustement au modèle théorique, la validité prédictive nous intéresse particulièrement, dans la perspective d'une utilisation de cet outil dans la conduite d'études consommateurs. Pour ce faire, nous avons comparé les valeurs du R^2 estimées pour les différentes relations postulées dans notre modèle théorique. Dans la *condition DS*, 60% de la variance des intentions de fidélité et 54% de la variance des intentions de bouche-à-oreille sont expliquées par la qualité perçue et la satisfaction globale. Le pouvoir prédictif du modèle passe à 59% de variance expliquée pour les intentions de fidélité et à 61% de variance expliquée pour les intentions de bouche-à-oreille dans la *condition ESG*. Ainsi, l'ESG semble proposer une validité prédictive supérieure aux traditionnels DS pour la variable d'intention bouche-à-oreille (tableau 26).

Construits	R^2	
	Condition SD	*Condition ESG*
Satisfaction globale	0,53	0,44
Intentions de fidélité	*0,60*	*0,59*
Intentions de bouche-à-oreille	*0,54*	*0,61*

Tableau 27. Pourcentages de variance expliquée dans chaque ensemble de données

Globalement, le modèle théorique testé est corroboré par les données collectées dans les deux conditions puisque toutes les relations sont positives et significatives ($p<0,05$). Nous

devons toutefois relativiser cette conclusion compte tenu des valeurs des RMSEA de nos modèles de mesure. La figure 28 présente les coefficients standardisés des relations entre les construits pour les deux modèles testés.

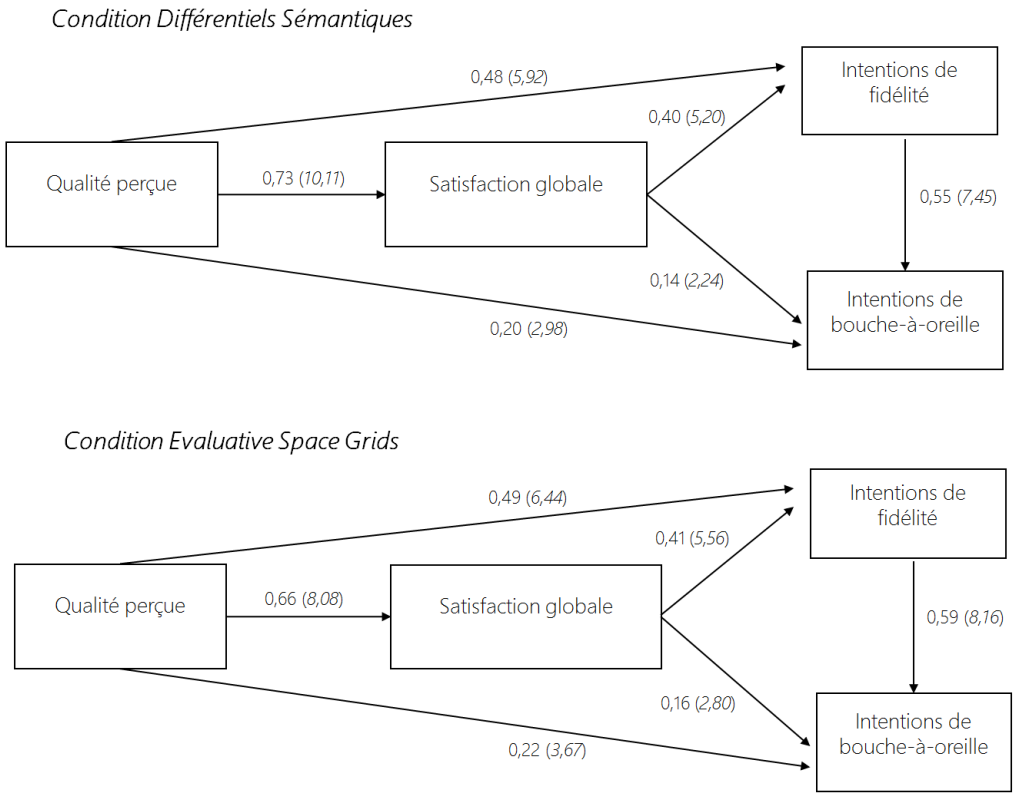


Figure 28. Bêta standardisés et valeurs de t (entre parenthèses) pour le modèle structurel

Afin de confirmer ces résultats, nous avons adopté la démarche d'analyse des effets indirects de médiation préconisée par Preacher et Hayes (2008). La méthode du *bootstrap* est une méthode très robuste pour le test des effets indirects (Zhao, Lynch et Chen, 2010). Elle permet de dépasser les limites de la démarche de Baron et Kenny (1986) grâce à l'utilisation des intervalles de confiance (Edwards et Lambert, 2007 ; MacKinnon, Lockwood et Williams, 2004). Cette méthode repose sur l'utilisation d'une macro SPSS qui allie le test de Sobel à une démarche en pas à pas qui permet de tester tous les effets indirects tout en contrôlant pour les autres variables du modèle. Les analyses se basent sur 5000 réplications générées par la méthode du *bootstrap*. Grâce à cette méthode, nous avons testé le rôle médiateur de la satisfaction dans la relation entre la qualité perçue et les intentions de comportement.

Les résultats sont similaires pour les deux ensembles de données. Tout d'abord, la qualité perçue a un effet positif et significatif sur les intentions de fidélité (DS : a=,57 ; t=6,42 ; p<,000 / ESG : a=,66 ; t=6,83 ; p<,000). En contrôlant par la qualité perçue, l'effet de la satisfaction sur les intentions de fidélité est positif et significatif (DS : b=,32 ; t=5,79 ; p <,000 / ESG : b=,44 ; t=6,40 ; p <,000). De plus, l'effet total de la qualité perçue sur les intentions de fidélité est significatif (DS : c=,98 ; t=9,87 ; p <,000 / ESG : c=,80 ; t=8,71 ; p <,000). Enfin, l'effet indirect de la qualité perçue sur les intentions de fidélité, à travers la satisfaction, est positif et significatif (DS : a x b =,32 ; z=4,98 / ESG : a x b =,36 ; z=5,13), avec un intervalle de confiance de 95% excluant 0 (DS : ,20 à ,47 / ESG : ,23 à ,53). Ces résultats indiquent une médiation complémentaire dans laquelle l'effet de la qualité perçue sur l'intention de fidélité passe à la fois par une voie directe que par l'intermédiaire de la satisfaction globale (Zhao, Lynch et Chen, 2010) (tableau 27 et 28).

Condition DS			Test des effets directs					Test des effets indirects				
Relations			Bêta	t	Borne inf. IC	Borne sup IC	Sig.	Bêta	z	Borne inf. IC	Borne sup IC	Sig.
Qualité perçue	→	Fidélité	0.57	6,41	0.39	0.75	0.000	0.32	4,98	0.20	0.47	0.000
Satisfac-tion	→	Fidélité	0.32	5,79	0.21	0.43	0.000					
Qualité perçue	→	Satisfac-tion	0.98	9,87	0.79	1.18	0.000					
Qualité perçue	→	Bouche-à-oreille	0.69	7,40	0.50	0.87	0.000	0.31	4,71	0.16	0.49	0.000
Satisfac-tion	→	Bouche-à-oreille	0.31	5,38	0.20	0.43	0.000					
Qualité perçue	→	Satisfac-tion	0.98	9,87	0.79	1.18	0.000					

Borne inf./sup. IC : borne inférieure/supérieure de l'intervalle de confiance ; Sig. : significativité

Tableau 28. Bêta standardisés et valeurs de t pour les effets directs et indirects de médiation suivant la méthode du *bootstrap* dans la *condition DS*

Il en va de même pour les intentions de bouche-à-oreille. La qualité perçue a un effet positif et significatif sur les intentions de bouche-à-oreille (DS : a=,69 ; t=7,40 ; p<,000 / ESG : a=,66 ; t=7,52 ; p<,000). En contrôlant par la qualité perçue, l'effet de la satisfaction sur les intentions de bouche-à-oreille est positif et significatif (DS : b=,31 ; t=5,40 ; p <,000 / ESG : b=,42 ; t=6,60 ; p <,000). De plus, l'effet total de la qualité perçue sur les intentions de bouche-à-oreille est significatif (DS : c=,98 ; t=9,87 ; p <,000 / ESG : c=,80 ; t=8,71 ; p

<,000). Enfin, l'effet indirect de la qualité perçue sur les intentions de bouche-à-oreille, à travers la satisfaction, est positif et significatif (DS : a x b =,31 ; z=4,71 / ESG : a x b =,34 ; z=5,32), avec un intervalle de confiance de 95% excluant 0 (DS : ,16 à ,49 / ESG : ,21 à ,51). Ces résultats indiquent également une médiation complémentaire (Zhao, Lynch et Chen, 2010) (tableau 27 et 28).

Condition ESG			Test des effets directs					Test des effets indirects				
Relations			Bêta	t	Borne inf. IC	Borne sup IC	Sig.	Bêta	z	Borne inf. IC	Borne sup IC	Sig.
Qualité perçue	→	Fidélité	0.66	6,83	0.47	0.85	**0.000**	0.35	5,13	0.23	0.53	**0.000**
Satisfaction	→	Fidélité	**0.44**	6,40	0.31	0.58	**0.000**					
Qualité perçue	→	Satisfac-tion	0.80	8,71	0.62	0.99	**0.000**					
Qualité perçue	→	Bouche-à-oreille	0.66	7,52	0.49	0.84	**0.000**	0.34	5,32	0.21	0.51	**0.000**
Satisfac-tion	→	Bouche-à-oreille	**0.42**	6,60	0.29	0.55	**0.000**					
Qualité perçue	→	Satisfac-tion	0.80	8,71	0.62	0.99	**0.000**					

Borne inf./sup. IC : borne inférieure/supérieure de l'intervalle de confiance ; Sig. : significativité

Tableau 29. Bêta standardisés et valeurs de t pour les effets directs et indirects de médiation suivant la méthode du *bootstrap* dans la *condition ESG*

La méthode du *bootstrap* nous permet de confirmer les liens directs et indirects de médiation postulés dans le modèle théorique étudié. Plus précisément, l'hypothèse centrale de médiation partielle de la satisfaction dans la relation entre la qualité perçue et les intentions de fidélité d'une part, ainsi qu'entre la qualité perçue et les intentions de bouche-à-oreille d'autre part, est validée.

Section 4. Conclusion

Cette étude a permis de valider l'*Evaluative Space Grid* comme outil de mesure alternatif aux différentiels sémantiques pour mesurer la satisfaction client. Plus précisément, au sein du modèle théorique testé, la grille de Larsen présente des performances globalement similaires aux différentiels sémantiques en termes de fiabilité, de validité convergente et de la validité discriminante.

Cependant, les résultats de cette étude ne nous permettent pas de valider l'hypothèse H1 qui anticipait un meilleur ajustement du modèle de mesure lorsque la satisfaction était mesurée avec des grilles de Larsen plutôt qu'avec des différentiels sémantiques. En effet, les indicateurs d'ajustement étant pour ainsi dire similaires dans les deux conditions expérimentales (DS et ESG), il nous est impossible de rejeter H0.

La deuxième hypothèse stipulait une meilleure validité prédictive dans la condition de mesure utilisant des grilles de Larsen. Ce n'est le cas que pour les intentions de bouche-à-oreille (54% de variance expliquée dans la *condition DS* contre 61% dans la *condition ESG*). Ce résultat prend du sens lorsque l'on examine attentivement les variables dépendantes étudiées. Pour rappel, les intentions de fidélité et de bouche-à-oreille étaient mesurées sur des échelles de Likert en 7 points indiquant un degré d'accord avec des propositions concernant des intentions de comportement. Cependant, il est intéressant de s'interroger sur les comportements associés à de faibles notes sur ces deux échelles. Un faible score sur une échelle d'intentions de fidélité est supposé se traduire dans les faits par une absence de réachat ou par le choix d'une marque concurrente. C'est un peu plus subtil lorsqu'un répondant attribue un faible score sur une échelle d'intentions de bouche-à-oreille positif. Cela peut vouloir dire qu'il n'a pas l'intention de parler de la marque en question, qu'il a l'intention de parler d'autres marques ou bien qu'il a l'intention de dire des choses négatives sur la marque évaluée. Nous pensons que l'*Evaluative Space Grid*, en distinguant le niveau de satisfaction et le niveau d'insatisfaction a permis une meilleure appréhension des variations de la variable « bouche-à-oreille » qui évolue elle aussi vers le positif et vers le négatif.

Il apparaît cependant nécessaire de soulever certaines limites pour tenter de comprendre les résultats mitigés de cette étude.

Section 5. Limites

Cette étude constituant la première utilisation de l'*Evaluative Space Grid* en tant que mesure de la satisfaction globale, nous avons été amenés à proposer un index dont le but était de saisir la part d'ambivalence contenue dans l'évaluation. Malgré les arguments théoriques apportés pour justifier nos choix, il est indiscutable qu'un codage différent fournirait des résultats différents.

Pour contourner cette difficulté, une approche alternative mériterait d'être creusée. La littérature portant sur les index de mesure soulève parfois la question du caractère intervalle (supposant l'égalité des distances entre chaque point de l'échelle) ou catégoriel, correspondant à des catégories distinctes et éventuellement ordonnables des outils (Thornton, 2011). On peut imaginer que les quatre types de réactions évaluatives décrites dans la partie théorique correspondent simplement à quatre catégories différentes. Dans cette perspective, il ne serait pas possible de créer un index de satisfaction globale permettant d'ordonner les répondants du plus insatisfait au plus satisfait en passant par l'ambivalent et l'indifférent et ce suivant des intervalles réguliers. Cette hypothèse fait l'objet d'un développement plus particulier dans un article de recherche réalisé en co-écriture avec Svein Ottar Olsen et Ana Alina Tudoran (encadré 1).

Encadré 1. Détails sur l'article « *Are you Rather Satisfied or Dissatisfied? A Comparative Evaluation of Traditional and Alternative Scales for Measuring Overall Satisfaction* », projet développé avec Svein Ottar Olsen et Ana Alina Tudoran

Dans cet article, en plus du test du modèle d'équations structurelles, une approche bayesienne est développée. L'analyse bayesienne présente au moins deux avantages par rapport à la modélisation par analyses structurelles. Tout d'abord, en utilisant le réseau bayésien on peut estimer les relations entre des variables latentes mesurées sur des échelles nominales ou ordinales. Plus précisément, cette technique permet la représentation de distributions de probabilités conditionnelles entre plusieurs variables catégorielles. Elle permet donc de traiter les données issues de l'*Evaluative Space Grid* en les associant à différentes catégories (satisfait, insatisfait, indifférent ou ambivalent). Les réseaux bayésiens permettent également d'envisager des relations non linéaires entre les variables (Neapolitan, 2004). En effet, nous avons postulé dans l'étude 3 de la thèse qu'il existait une relation linéaire entre la satisfaction globale et les variables dépendantes. Il est toutefois possible que la prise en compte de la part d'ambivalence contenue dans la satisfaction globale vienne altérer la linéarité de la relation avec les variables dépendantes. Les résultats obtenus en traitant les données issues des deux formats alternatifs de mesure (DS et ESG) comme des variables catégorielles, soutiennent les conclusions de l'étude 3 de la thèse : les deux outils de mesure offrent des validités convergente, discriminante, de construit et prédictive relativement similaires.

Par ailleurs, la modélisation par analyses structurelles n'a pas pour premier objectif la comparaison de modèle de mesure, mais plutôt l'ajustement d'un modèle de mesure à un modèle théorique. En ce sens, les résultats de l'étude 3 ne sont pas tout à fait conformes à nos hypothèses puisque nous avions anticipé qu'en améliorant en particulier la validité de construit, le modèle de mesure utilisant des *Evaluative Space Grids* s'ajusterait mieux au modèle théorique que celui utilisant des différentiels sémantiques. Si la taille réduite des

échantillons dans les deux conditions permet sans doute d'expliquer la faiblesse des indices RMSEA, nous pensons que la difficulté d'utilisation des grilles de Larsen permet peut-être d'expliquer pourquoi les résultats sont légèrement moins bons dans la *condition ESG*. En effet, nous manquons d'élément concernant l'interprétation faite par les répondants des différentes cellules de la grille.

***

Dans le but de mieux comprendre comment les individus s'approprient ce nouvel outil, le chapitre 5 de la thèse est donc consacré à des entretiens d'experts (chercheurs et professionnels des études) qui se sont penchés sur la compréhension de la grille.

Chapitre 1	
Revue de la littérature	Le concept d'ambivalence La mesure de l'ambivalence
Etude préliminaire	La mesure des évaluations globales de consommateurs dans le e-commerce : pratiques et perception des internautes

↓

Chapitre 2	
Cadre conceptuel et hypothèses	Présentation de *l'Evaluative Space Grid* Résultats attendus

↓

Chapitre 3	
Etude 1	Validité convergente de *l'Evaluative Space Grid* avec les paires de continuums unipolaires
Etude 2	Validité convergente de *l'Evaluative Space Grid* avec les différentiels sémantiques, les discours des consommateurs et une mesure de l'ambivalence subjective

↓

Chapitre 4	
Etude 3	Test de *l'Evaluative Space Grid* comme outil de mesure de la satisfaction client

↓

Chapitre 5	
Etude 4	Etude qualitative après d'un échantillon d'experts : - qu'en pensent les chercheurs ? - qu'en pensent les professionnels ?

↓

Chapitre 6	
Etude 5	Place de l'incertitude dans les évaluations effectuées avec *l'Evaluative Space Grid*

– Chapitre 5 –

Processus d'appropriation de l'*Evaluative Space Grid* par les répondants : l'avis des experts

Comment les individus utilisent-ils l'*Evaluative Space Grid* pour effectuer des évaluations ? Nous souhaitons approfondir cette question pour comprendre les résultats mitigés du chapitre précédent qui se conclut sur l'équivalence psychométrique entre l'*Evaluative Space Grid* et les différentiels sémantiques. L'hypothèse d'un taux d'erreur de réponses supérieur avec cet outil alternatif nécessite d'être étayée empiriquement. Par ailleurs, nous manquons d'éléments quant au sens donné par les répondants aux différentes modalités de réponses de la grille de Larsen.

Pour approfondir la compréhension des processus d'appropriation de l'*Evaluative Space Grid* par les répondants, nous avons interrogé deux profils d'experts des questions de mesure : huit chercheurs en sciences sociales (Section 1) et cinq professionnels des études et des avis de consommateurs en ligne (Section 2).

Section 1. Qu'en pensent les chercheurs ?

1. Présentation de l'étude et méthode

1.1. Nombre d'entretiens et choix de l'échantillon

L'échantillon de participants de la première phase de cette étude était composé de chercheurs en marketing, en management et en sociologie. Nous avons choisi d'interroger des experts ayant déjà été confrontés à des problématiques de mesure car nos questions étaient techniques et demandaient de bonnes capacités d'abstraction (Richardson, 2000). Le tableau 29 décrit le profil des personnes interviewées. Les prénoms de ces personnes ont été modifiés pour des raisons de confidentialité.

Participants	Sexe, âge et profession
Nathalie	Femme, 34 ans, enseignant-chercheur en marketing, ancienne chef de produit
Elodie	Femme, 32 ans, enseignant-chercheur en management
Lucie	Femme, 32 ans, enseignant-chercheur en marketing
Paul	Homme, 37 ans, enseignant-chercheur en marketing
Nabih	Homme, 35 ans, enseignant-chercheur en management
Louisa	Femme, 26 ans, doctorante en sociologie
Najate	Femme, 35 ans, enseignant-chercheur en management
Anna	Femme, 24 ans, doctorante en stratégie

Tableau 30. Profil des chercheurs interrogés

Compte tenu de nos objectifs, le nombre d'entretiens à réaliser a été fixé suivant le critère de saturation sémantique ou saturation théorique. Elle est atteinte lorsque l'ajout d'un entretien supplémentaire n'apporte plus de nouvelles informations par rapport aux précédents entretiens (Miles et Huberman, 1991). Huit interviews ont permis de satisfaire ce critère.

1.2. Déroulement des entretiens

Sur les huit entretiens semi-directifs réalisés, quatre se sont déroulées à l'Université Paris-Dauphine et quatre autres à l'Institut Supérieur de Gestion de Paris. Ils ont duré entre 35 et 75 minutes. Un questionnaire en ligne servait de support à l'entretien. De manière à créer

une situation similaire à celle de la passation d'une enquête en ligne, les rencontres se sont déroulées dans les bureaux des personnes interrogées, face à leur ordinateur. L'objectif était de faire expliciter par les participants les processus de réponse à des questions sous format de grilles de Larsen. Plus précisément, les entretiens se divisaient en quatre parties :

- Phase 1 : les participants étaient invités à répondre à des questions portant sur leur satisfaction à l'égard de trois services ;

- Phase 2 : ils devaient ensuite se placer dans la peau d'un chargé d'études et essayer d'analyser une série de réponses rapportées sur des grilles de Larsen par des consommateurs fictifs ;

- Phase 3 : puis il leur était demandé d'indiquer comment ils envisagent la transposition des réponses issues d'*Evaluative Space Grids* sur des différentiels sémantiques ;

- Phase 4 : enfin, nous leur demandions d'annoter une grille imprimée sur une feuille en les invitant à synthétiser la signification qu'ils attribuaient à chaque zone de la grille[10].

Dans le prolongement de l'étude précédente, toutes les grilles présentées dans le cadre de ces entretiens étaient appliquées à la mesure de la satisfaction globale (avec « satisfaction » en abscisse et « insatisfaction » en ordonnée). La méthode des protocoles verbaux (Kuusela et Paul, 2000) était utilisée dans les phases 1 et 3 de l'entretien, en incitant les participants à verbaliser l'ensemble de leurs pensées. La technique des protocoles verbaux a été mise au point au début des années 80 par les psycholinguistes cognitivistes américains Hayes et Flower qui se sont intéressés aux processus cognitifs à l'œuvre dans l'écriture. L'objectif est d'accéder à ce qui se passe « dans la tête » du sujet au moment où il réalise une tâche : à quoi il pense, pourquoi il effectue cette tâche de cette manière, etc (Forget, 2013).

Nous avons adopté une technique d'interview semi-directive dans le but de laisser aux participants une multitude de façons de répondre aux objectifs assignés par l'entretien. L'intégralité des discussions a été enregistrée. Le guide d'entretien est présenté en annexe 4.

[10] *Cette quatrième phase a été initiée spontanément par la troisième personne interrogée. Nous avons alors décidé de l'intégrer à notre protocole d'entretien, expliquant pourquoi seuls six des huit participants se sont livrés à cet exercice.*

2. Analyses

Les discours obtenus ont été intégralement retranscrits puis analysés à partir d'une analyse de contenu catégorielle thématique effectuée sous Excel. Nous avons opté pour une procédure suivant une analyse thématique ad hoc « ouverte » (Ghiglione et Matalon, 1978) dans la mesure où nous ne disposons d'aucune littérature sur l'utilisation de la grille pour définir des thèmes préétablis. Après lecture flottante, chaque retranscription a fait l'objet d'un premier codage général sur quatre grandes catégories. Ensuite, un second codage fait émerger des sous-catégories de l'arbre de codage. Chaque information est affectée à une seule rubrique. Toutefois, la troisième phase des entretiens, consacrée à la transposition de notes issues de la grille sur des différentiels sémantiques n'est pas intégrée à l'analyse. En effet, les verbatims de cette partie sont peu développés dans la mesure où la plupart des répondants se sont livrés à des calculs à voix haute, rendant l'analyse difficile.

3. Résultats

La figure 19 présente l'arbre de codage issu de l'analyse.

3.1. Prise en main de l'*Evaluative Space Grid*

Compréhension de l'outil

Les réactions des participants ont laissé transparaitre deux étapes d'appropriation de la grille de Larsen. Dans un premier temps c'est l'étonnement qui primait face à cette « *double échelle* » (Elodie). Les participants se sont interrogés sur la pertinence d'une mesure distincte pour la satisfaction et l'insatisfaction, puisque selon eux, « *c'est deux mots similaires, enfin c'est les opposés* » (Louisa). C'est ce qu'illustrait également cette réaction de Nathalie : « *spontanément quand tu vois satisfaction / insatisfaction, tu as l'impression que c'est deux variables opposées* », ou bien les propos de Lucie qui s'est interrogée : « *Est-ce qu'on peut être extrêmement insatisfait et extrêmement satisfait en même temps ?* ». Anna est restée plus descriptive : « *la satisfaction et l'insatisfaction deviennent deux éléments séparés, non plus deux opposés d'un même continuum* ». Mais, très rapidement ils définissaient des conditions dans lesquelles une telle mesure était envisageable. Nathalie a souligné que « *ça peut aussi être des critères différents* » et Nabih a conclut que ça permet de produire « *un jugement global par rapport à plusieurs critères* ».

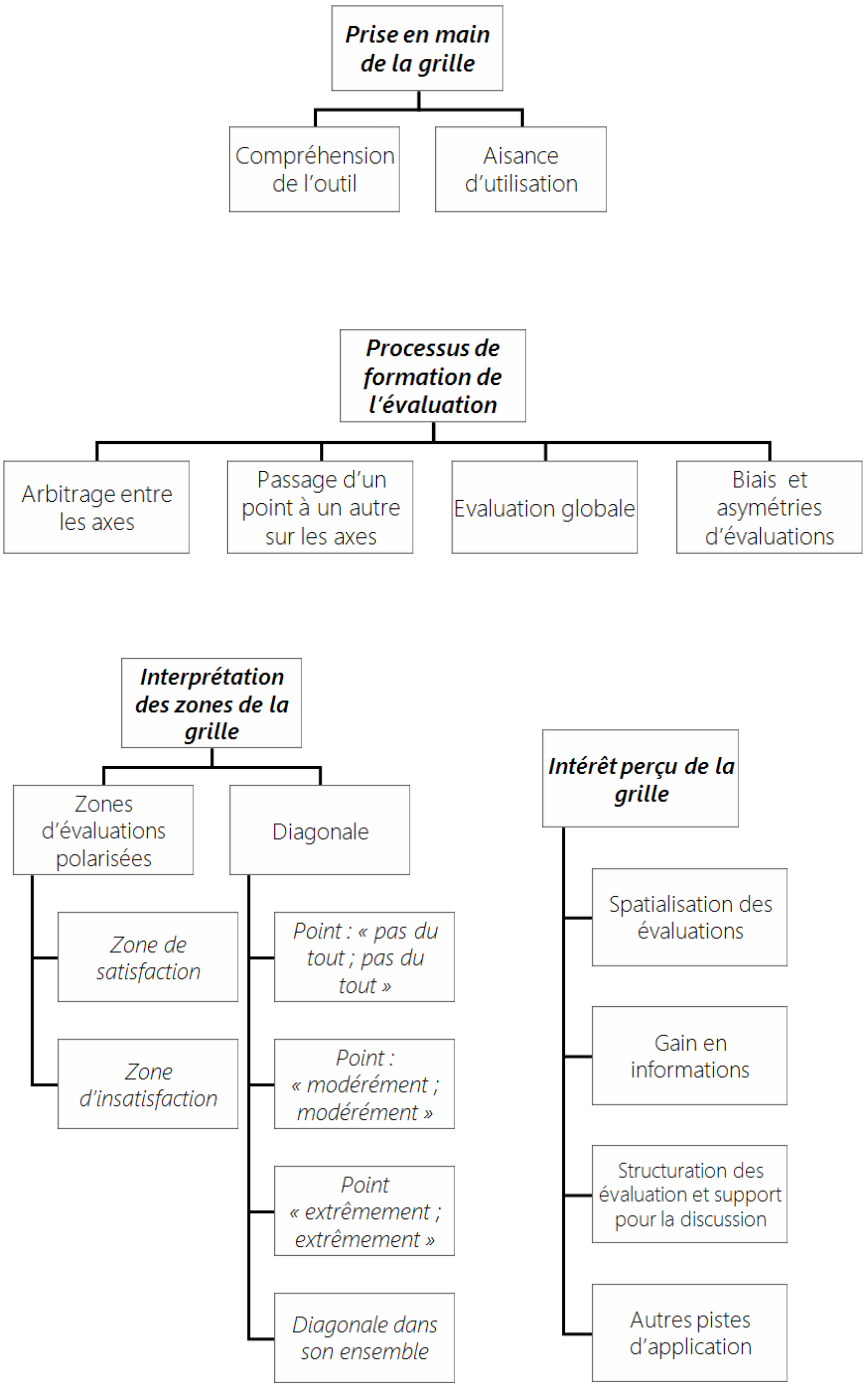


Figure 29. Arbre de codage des entretiens semi-directifs

Aisance d'utilisation

Certains verbatims traduisent clairement des difficultés d'utilisation. Au départ, Elodie a été un peu gênée par le format de l'outil : « *J'ai du mal à me projeter dans une double échelle*

en fait, c'est le premier truc qui me… qui me perturbe ». Lucie trouvait quant à elle « *qu'on n'a pas un rapport spontané avec* » et Nathalie qu' « *il y a un petit coût d'entrée* ».

Plus précisément, les participants pensent que « *sur cette matrice en fait il y a des zones qui sont vraiment claires et de zones qui sont plus difficiles* » (Lucie), autrement dit « *il y a des réponses plus cohérentes que d'autres* » (Nabih). Lucie a tenté d'expliquer ce phénomène par le fait « *qu'on est vraiment formaté avec ce continuum entre satisfaction et insatisfaction […]. Donc ça oblige à une gymnastique intellectuelle qui pour moi au début effectivement, me déroutait, mais au final enfin je trouve qu'au bout d'un moment en utilisant, on se fait bien à l'outil* ».

Quelques erreurs d'utilisation durant les entretiens traduisent concrètement ces difficultés. Après avoir inversé le sens de lecture des axes Nathalie s'est exclamée : *« non ben non ben non ! non ! pas du tout ! […] Mauvaise interprétation !* ». Même problème pour Anna qui a constaté « *C'est plus une… mauvaise interprétation […] il faut être attentif* ». Pour faciliter ce « *petit apprentissage léger à faire* », Lucie nous a conseillé d'ajouter « *quelques phrases introductives* » : « *Moi j'aurais besoin qu'on me dise au début du genre, imaginez que vous avez plusieurs sources… […]. Dans votre évaluation, vous avez plusieurs éléments… Et j'aurais besoin qu'on me dise par exemple vous avez… […]. vous avez trois éléments de satisfaction deux éléments d'insatisfaction ça fait ça sur la matrice. Ou alors vous en avez deux et deux, mais il y en a deux qui sont plus importants ça fait ça sur la matrice ».*

Cette dernière remarque nous amène naturellement à développer la manière dont les individus restituent leurs réactions sur l'*Evaluative Space Grid*.

3.2. Processus de formation de l'évaluation

Les participants ont verbalisé plusieurs étapes pour expliciter le passage de l'évaluation formée dans leur esprit à la note donnée sur la grille de Larsen. Elodie a décri le processus de la manière suivante : « *je pense que je fonctionne en quatre étapes. D'abord je vois la satisfaction, d'accord ? Donc j'aurais tendance à me dire… pourquoi je suis satisfaite ? Je me pose les arguments. Après je vais me dire pourquoi je suis insatisfaite je me pose les arguments. Et, une fois que j'ai posé les deux, j'essaie de regarder comment je suis satisfaite… J'essaie de faire la balance entre les deux, mais en même temps en me plaçant d'abord sur satisfaction et après sur insatisfaction, tu vois ce que je veux dire ? C'est-à-dire que je pense que l'un agit sur l'autre.* »

Arbitrage entre l'axe de la satisfaction et l'axe de l'insatisfaction

Dans la mesure où la grille de Larsen propose de déployer sur deux dimensions des réactions qui sont habituellement concentrées sur un continuum, il est important de comprendre sur quels critères les répondants décident de qualifier une réaction de forte satisfaction (resp. insatisfaction) plutôt que de faible insatisfaction (resp. satisfaction). Louisa a clairement caractérisé ce processus « *d'arbitrage* ». Pour Elodie, « *la satisfaction globalement se porte sur certains aspects du service et l'insatisfaction se porte sur d'autres* », elle a également parlé de « *deux caractéristiques différentes* » ou de « *critères différents pour évaluer la satisfaction et l'insatisfaction* ». Nathalie a expliqué au sujet de son opérateur de téléphonie mobile : *« je suis un peu satisfaite et un peu insatisfaite, mais pas sur les mêmes attributs* ». Cette idée a été reprise par Najate de la manière suivante : « *il y a des points positifs qui font que je suis satisfaite et d'autres négatifs qui font que je sais qu'il y a une marge d'amélioration* ». Lucie a précisé quant à elle qu'il y a « *un axe où [elle est] satisfaite et un axe où* [elle est] *insatisfaite* », alors que Paul parlait « *d'aspects* ». Pour Najate, il s'agissait de « *facteurs différents* » et Elodie a exprimé cela en termes « *d'éléments du service* ». Au final Lucie a verbalisé ainsi l'évaluation qu'elle faisait de son shampooing : « *je suis extrêmement satisfaite de… de ses propriétés lavantes, c'est doux, machin, par contre je suis insatisfaite, "beaucoup" on va dire, par l'odeur* ».

De manière intéressante, plusieurs participants ont évoqué un effet de « *vases communicants* » (Najate). Paul décrivait ce phénomène dans les termes suivants : « *Il y a quand même une balance qui se fait je pense… entre la satisfaction et l'insatisfaction […] Quand tu as du "peu" à un endroit t'as… t'as forcément…au moins du "modérément" de l'autre côté […] dès lors que tu as trop avancé sur un axe, tu es obligé de reculer sur l'autre* ». Quand l'avis est tranché, la communication entre les deux axes est telle que toute la note se porte sur l'un des deux axes comme le notait Paul « *un tel désagrément que tout est… le négatif a absorbé tout l'effet…* »

Cette idée de « *vases communicants* » entre la satisfaction et l'insatisfaction est tout à fait cohérente avec l'approche bidimensionnelle dans laquelle nous nous inscrivons. En effet, Cacioppo et ses collègues (Cacioppo et Berntson, 1994 ; Cacioppo & al., 2011 ; Cacioppo, Gardner et Berntson, 1997) proposent une distinction partielle entre les systèmes d'évaluation de la positivité et de la négativité. Dans ce modèle, les effets réciproques entre les évaluations positives et négatives existent mais ne sont pas systématiques comme semblaient le percevoir les participants aux entretiens semi-directifs.

Passage d'un point à l'autre sur les axes

Les participants ont justifié le passage d'un point à un autre le long des deux axes de différentes manières. Il peut tout d'abord s'agir d'une question d'intensité : « *mes réactions sont beaucoup plus fortes* », « *les critères de satisfaction sont forts* », « *j'ai eu des critères de satisfaction, mais j'hésite quant à leur force, leur intensité* » (Nathalie). L'importance accordée au critère est également évoquée : « *ça dépendra vraiment beaucoup de la pondération aussi, de l'importance que j'accorde à ces différents critères* » (Najate). Pour Nabih le répondant va passer en revue « *tous les critères […] qu'il considère comme importants* ». Cette idée de force, d'importance ou d'intensité se retrouvait également dans les propos de Louisa lorsqu'elle a justifié le passage de « beaucoup » à « extrêmement » sur l'axe de la satisfaction : « *Moi je trouve que quand tu vas sur extrêmement, entre beaucoup et extrêmement, c'est que extrêmement tu… vraiment tu… ça t'emballe quoi. Alors que beaucoup, tu aimes bien. Tu vois c'est un peu un truc… Et puis je pense qu'il y a aussi un truc… Tu vois par exemple extrêmement, genre c'est ce que tu aimes manger quoi, par-dessus tout, c'est ce que tu aimes manger. Alors que beaucoup, je dirais… c'est que… tu aimes bien. De temps en temps.* »

Un deuxième critère, celui du nombre d'éléments évalués sur chaque axe, a été invoqué par Elodie sous la forme suivante : « *je ne sais pas si on peut quantifier en beaucoup pas beaucoup…* ». Cette idée a été reprise par Louisa qui nous a expliqué que pour "modérément" elle « *pourrait trouver deux ou trois trucs et donc effectivement… qui s'accumulent un peu et qui, prennent le pas sur [son] jugement* ». Cette idée est également apparue dans le discours de Najate nous expliquant qu'elle avait « *une ou deux raisons d'être insatisfaite* » pour justifier une faible note sur l'axe de l'insatisfaction.

Elodie et Lucie percevaient clairement que le passage d'un point à un autre le long d'un axe pouvait s'expliquer de deux manières : « *avec des facteurs plus nombreux ou plus marqués* » (Elodie). Lucie a verbalisé les choses de la manière suivante « *soit il y a plus d'éléments de satisfaction que d'insatisfaction. Par exemple il y a quatre éléments de satisfaction et il y a que trois éléments d'insatisfaction, voilà. Soit il y a trois éléments de satisfaction et trois éléments d'insatisfaction mais les éléments de satisfaction sont… d'autres éléments que… Et ils sont plus importants* ». Elle a intégré d'ailleurs spontanément les critères de l'importance et du nombre d'élément lorsqu'elle a analysé une grille de Larsen remplie par un répondant fictif lors de la deuxième phase de l'entretien : « *Il y a peu*

d'éléments qui font qu'elle est insatisfaite, donc il y en a peut-être juste un petit mais c'est pas très important, donc ça pourrait être, je ne sais pas moi, la musique par exemple ».

Il semble donc qu'il y ait une réelle perméabilité entre les critères de l'importance et du nombre d'éléments pris en compte. Cela nous invite à penser que les deux aspects peuvent être pris en considération en fonction des situations ou des individus, ce qui n'est pas étonnant dans le cadre d'une évaluation globale.

Evaluation globale

Pour la plupart des répondants, l'évaluation globale résulte « *d'une espèce d'équilibre entre ce qui [leur] plaît et ce qui ne [leur] plaît pas* » (Paul). Autrement dit, « *on doit faire une balance* » (Elodie) et « *pondérer tout ça* » (Elodie). S'ils ne sont pas rentrés dans l'analyse des différents poids accordés à « *ce qui leur plaît* » et « *ce qui ne leur plait pas* », les participants ont détaillé les étapes du processus de formation de l'évaluation. Louisa l'a décrit dans les termes suivants : « *Moi je commence à regarder la satisfaction, je ne sais pas pourquoi je commence par ça, à me dire pourquoi je suis satisfaite* ». Au-delà de la simple chronologie évaluative, ce verbatim nous éclaire sur la localisation du point d'ancrage au sein de la grille. En effet, selon le paradigme de la disconfirmation des attentes, la satisfaction résulte d'un processus de comparaison entre l'expérience subjective du consommateur, ou performance perçue, et les attentes qu'il a constituées au fil des expériences (Oliver, 1980 ; 1997).

Nous pensons que le format de l'outil employé pour effectuer l'évaluation peut avoir un impact sur le point d'ancrage du répondant. Ainsi, pour Louisa, c'était le niveau de satisfaction qui jouait le rôle de point d'ancrage dans son évaluation. On retrouvait la même logique dans les propos de Nabih : *« J'ai l'impression que, quand j'évalue un truc, c'est en termes de satisfaction qui va de zéro jusqu'à plus l'infini, et je ne parle d'insatisfaction que quand il se passe des trucs qui me rendent réellement insatisfait. Donc des trucs assez forts* ». Anna semblait également fonctionner de cette manière : « *finalement j'ai un niveau, on va dire il y a un niveau d'exigence standard, je m'attends qu'à minima, mon opérateur mobile fasse euh, fasse telle ou telle chose […] et à partir de ce standard, je vais plutôt mettre des points négatifs à chaque fois qu'il y a un truc qui ne va pas* ». Dans d'autres cas, Anna prend l'insatisfaction pour point d'ancrage : « *pour le coiffeur… je dirais que je partirais plutôt d'une logique de critères d'insatisfaction puisque… voilà. Pour moi le coiffeur c'est plutôt un risque de… que ça va être cher, il y a un risque que la coupe soit ratée, que le salon*

ne soit pas agréable. […] Donc c'est vrai que par exemple, si je ressors… avec une coupe qui m'a beaucoup plu, ça va vraiment produire de la satisfaction plutôt que de neutraliser mon insatisfaction ».

Ces différents extraits illustrent le fait que la grille s'adapte très bien à une logique évaluative basée sur la disconfirmation des attentes. En fonction de leurs expériences passées, les participants s'attendent plutôt à être satisfaits ou insatisfaits ; ils ancrent leur évaluation sur l'axe en question, puis ils ajustent la note globale en se positionnant sur le deuxième axe.

Les participants nous ont également apporté des éléments sur la nature du résultat issu de ce processus évaluatif. Pour Paul, ce n'est pas directement la note donnée sur la grille qui relève d'une approche globale, mais plutôt les coordonnées de cette note sur l'axe de la satisfaction et l'axe de l'insatisfaction : « *c'était peut-être pas une vraie approche globale c'était peut-être une espèce de… d'approche globale de l'insatisfaction, d'approche globale de la satisfaction, tu vois, pas sur des attributs précis* ». Autrement dit, l'analyse holistique peut se faire séparément pour chaque axe.

Dans cette perspective, l'outil serait plutôt adapté à des personnes ayant un esprit analytique. Selon Lucie : « *c'est un peu un état d'esprit enfin… une… des gens qui ont des caractères différents… qui… La manière dont ils répondent à cette matrice… et il y a soit ceux qui ont un esprit plutôt holistique / global, ils n'ont pas trop… et ne sont pas trop analytiques ils ne séparent pas trop… les éléments de satisfaction et les éléments d'insatisfaction et à ce moment-là je m'imagine que cette personne serait un peu gênée par cette matrice et du coup… elle choisirait le même niveau pour les deux axes. Soit c'est une personne qui est… qui a un esprit un petit peu plus analytique et… Elle sépare bien les deux axes et il se trouve que… pour ces deux axes il y a le même niveau. J'imagine plus deux personnes différentes* ».

Pour conclure sur ce point, il semble que l'étape de globalisation se situe au niveau des axes « *et c'est la moyenne des deux qui fait que tu vas te positionner* » (Najate). Le processus de restitutions des réactions sur l'*Evaluative Space Grid* consiste donc à générer donc deux notes globales qui sont ensuite « mécaniquement » combinées au sein d'un point sur la grille.

Biais et asymétries d'évaluation

Les discussions sur la formation de l'évaluation globale ont également permis de mettre en évidence un certain nombre de biais possibles. Cela se traduit dans le discours des répondants par la description d'asymétries dans l'utilisation et l'interprétation qu'ils font des notes issues de la grille de Larsen.

Tout d'abord, le poids accordé aux deux axes ne semble pas tout à fait équivalent : « *je pense que l'égalité entre les deux penche quand même vers l'insatisfaction, c'est un peu le côté les… les pertes sont plus valorisées que le gain en fait* » (Paul). Cette idée est également apparue dans les propos de Nabih : « *Si je voulais réfléchir un peu plus, je me dirais que peut-être que l'insatisfaction marque beaucoup plus, et que le "beaucoup" de l'insatisfaction peut avoir plus de poids que "l'extrêmement" de la satisfaction. Ça dépend des personnes* ».

Cette asymétrie entre les axes est susceptible d'influencer l'interprétation que les participants font des différents points. Ainsi, sur l'axe de la satisfaction les notes faibles ont invité aux réflexions suivantes : « *Pour moi "pas du tout satisfait" c'est quand même… presque un peu négatif quoi* » (Lucie) ou encore : « *La faible satisfaction c'est déjà une forme d'insatisfaction* » (Nabih). Autrement dit : « *alors qu'on pourrait dire que ce "pas du tout", c'est l'absence et donc ça correspond un "zéro" mais, mais en même temps, dire heu… dire "pas du tout" c'est déjà prendre position […]. C'est un "1" en valeur absolue, mais oui c'est déjà une prise de position, effectivement j'avais pas réfléchi à cette dimension-là, mais "pas du tout de satisfaction", effectivement on peut dire que c'est déjà "-1"* » (Paul).

Au contraire, il est intéressant de noter que les faibles notes d'insatisfaction ne pas semblaient pas être associées à un « début de satisfaction ». La littérature en psychologie s'est intéressée à cette asymétrie. Elle souligne que, globalement, les individus accordent plus d'importance aux informations négatives qu'aux informations positives, en raison de leur rareté. En effet, la préférence naturelle des êtres humains pour le positif les conduit à effectuer une forme de sélectivité en faveur du positif dans leurs souvenirs, leur langage et leurs pensées (Matlin et Stang, 1978). C'est le principe de Pollyanna qui décrit cette préférence pour les aspects positifs ou favorables des choses. Entourés de positif, les individus réagissent plus fortement aux informations négatives, qu'ils jugent plus informatives (Baumeister & al., 2001 ; Rozin et Royzman, 2001).

Au-delà de ces asymétries largement étudiées dans la littérature, Paul considère que « *de toute façon, oui, les écarts entre les points ne sont pas les mêmes. Ça, ça me paraît une*

évidence. Ça me paraît une évidence que "beaucoup" et "extrêmement" ça tire vers le haut, et que "peu" ou "pas du tout" ça tire vers le bas et qu'en fait c'est resserré ». S'il est le seul à mettre des mots sur ce phénomène, la question de l'inégalité des distances entre les points ressort très clairement dans la troisième partie de l'entretien durant laquelle les participants devaient faire correspondre des notes issues de la grille de Larsen sur des différentiels sémantiques. L'exercice a révélé que certaines correspondances étaient difficiles à faire, sans doute justement à cause de cette distorsion entre les points. On retrouve ainsi les même biais que ceux constaté pour les échelles de Likert, considérées par convention comme des échelles intervalles, sans que l'hypothèse d'équidistance entre les échelons ne soit le plus souvent vérifiée (Chandon et Bartikowski, 2004).

Il est cependant peu aisé d'apporter sur ce point une réponse complète. En effet, c'est uniquement lorsqu'ils se placent dans la phase d'interprétation des réponses données par d'autres que les participants ont relevé ces asymétries. Nous pouvons en déduire que le chargé d'études cherchant à analyser les réponses issues de l'*Evaluative Space Grid* pourrait faire l'objet de biais similaires, mais il est difficile de savoir si ce type d'asymétries se forme également au moment où l'évaluateur tente de traduire son appréciation sur la grille. Pour mieux comprendre ce phénomène intéressons-nous au sens donné par les participants aux différentes zones de la grille.

3.3. Interprétation des différentes zones de l'*Evaluative Space Grid*

Les participants ont mis en évidence deux grandes zones évaluatives au sein de la grille de Larsen : la diagonale qui donne lieu à une importante discussion et les zones d'évaluations polarisées de chaque côté de celle-ci.

Les zones d'évaluations polarisées

La zone de satisfaction

On distingue tout d'abord une zone de satisfaction ou d'évaluations positives, même si sa délimitation peut varier légèrement selon les participants : *« les clients satisfaits sont toujours dans la moitié droite, donc si on trace un trait depuis "modérément", ils sont toujours euh… le meilleur cadran c'est le cadran en bas à droite. Ils sont pas insatisfaits et très satisfaits »* (Nathalie). Elodie a décri plus précisément cette zone : « *une zone de satisfaction assez franche dans le rectangle en bas à droite… De "modérément" à "extrêmement satisfait" et de "pas du tout" à "peu insatisfait"* ». Cette zone était un peu plus

large dans l'esprit de Louisa : « *Donc satisfaction "moyenne", "grande", "extrême" et insatisfaction "faible" ou "aucune" ben là vraiment c'est quelque chose qui me plaît quoi* ».

Zone d'insatisfaction

De l'autre côté de la grille *« lorsqu'on est dans beaucoup insatisfait jusqu'à modérément satisfait, c'est davantage des situations où on n'est pas satisfait »* (Najate). Louisa a défini plus précisément cette zone d'insatisfaction ou d'évaluations négatives dans les termes suivants : *« satisfaction "aucune", "faible" et insatisfaction "moyenne", "grande", "extrême", je dirais que là, ça ne me plaît pas. Que, en gros, le…les points négatifs prennent largement le pas sur les points positifs »*. Pour Paul, *« "extrêmement" d'insatisfaction et "pas du tout" de satisfaction je pense que c'est la… la colère… »*. Adoptant une analyse plus marketing, Nathalie nous a expliqué que *« là c'est la cata parfaite ! Puisque là t'es "très insatisfait" et t'es "pas du tout satisfait" donc là c'est des clients perdus. Donc c'est l'horreur »*.

La diagonale

« Tous les autres points restants, pour moi c'est une opinion assez… vraiment partagée on est un peu tiraillé, il y a à la fois du positif et du négatif et… c'est des éléments importants à chaque fois, qui sont… qui sont source de satisfaction ou… d'insatisfaction. Genre, il y a deux éléments importants dans la satisfaction, je sais pas moi, le prix et la couleur, imaginons le prix, la couleur, le service et il y a certains éléments qui sont satisfaits et un autre qui est… ou deux autres qui sont pas satisfaits » (Lucie). Cette « sorte de balance entre les facteurs de satisfaction et les facteurs d'insatisfaction » (Elodie) a été appréhendée par certains comme les *« dimensions les plus délicates à analyser »* (Elodie).

Voyons plus en détail comment ils ont interprété cette zone de *« partage »* (Paul).

Le point « pas du tout satisfait » / « pas du tout insatisfait »

Pour commencer *« il pose problème ce point »* (Nathalie). *« C'est possible ça ? »* nous a-t-elle demandé. Après quelques instants de réflexion elle conclut *« c'est pas d'avis en fait »*. C'est d'ailleurs l'interprétation qui est ressorti le plus fréquemment du discours des participants. Pour Lucie : *« Logiquement si tu n'es "pas du tout satisfait" et "pas du tout insatisfait" c'est censé être la réponse… "sans opinion" quoi. Logiquement, mathématiquement. Ça s'annule quoi »*. Paul a tenté de remonter aux causes d'une telle évaluation : *« c'est une case "ne sait pas" parce qu'il n'y a pas l'expérience du truc en fait.*

C'est un peu comme si… c'est un peu comme si tu me demandais d'évaluer un resto dans lequel je ne suis pas allé ». Cette même idée a été reprise par Louisa : *« tu t'en fous, tu ne sais pas. Parce que globalement tu n'as pas beaucoup d'arguments à mettre dans la balance, donc c'est soit, si tu n'as pas d'argument, c'est que tu ne sais pas, ou bien tu t'en fous. Soit tu n'as pas d'avis. Oui, c'est "pas d'avis". J'ai pas d'avis parce que je n'ai pas grand-chose à dire dessus, soit parce que tu ne connais pas, soit parce que ça ne t'intéresse pas »*. L'idée de l'absence de réaction revenait également dans le discours de Nabih : *« Mais encore une fois sur ça, "pas du tout/pas du tout", la seule explication que je peux trouver, c'est le fait de s'en foutre royalement, c'est de ne pas avoir de perception, d'émotion réelle »*.

Les participants ont associé cette absence de réaction à de l'indifférence : *« alors quand je ne suis pas impliquée émotionnellement je veux dire, je peux ressentir de l'indifférence, c'est-à-dire ça ne me fait ni chaud ni froid je m'en fous. Dès que ça m'implique un tout petit peu, j'aurais du mal à me situer là et à me dire je suis ni satisfaite, ni insatisfaite »* (Najate). Selon Elodie *« je sais pas si ça peut être assimilé à de l'indifférence par exemple, mais, là la personne semble ni satisfaite ni insatisfaite. Donc est-ce qu'elle a une opinion finalement sur le service de l'assurance habitation ? Là j'ai l'impression qu'il y a aucun critère qui génère un sentiment ou un avis quelconque en fait »*. Certains participants ont également mobilisé le concept de « neutralité » : *« "neutre" ça serait "pas du tout insatisfait/pas du tout satisfait" »* (Lucie).

Toujours à propos du point « pas du tout satisfait » / « pas du tout insatisfait », Nathalie a souligné une nuance intéressante : *« "indifférent" veut dire deux choses […] tu peux avoir des croyances fortes par exemple à l'égard de Domenech, mais que ces croyances n'aient pas de charge émotionnelle. Par exemple c'est un mauvais entraineur, pour moi "être un mauvais entraineur", ça n'a pas de valeur émotionnelle, pour mon mari, "être un mauvais entraîneur" c'est hyper important […] en fait ça veut aussi bien dire pas d'émotion, cette personne est indifférente, et ça veut aussi dire je ne pense rien »*. Elle a ainsi fait clairement apparaitre les deux types d'interprétations effectuées par les participants au sujet du point « pas du tout satisfait » / « pas du tout insatisfait » : l'absence de réaction et l'absence d'avis.

Certains participants sont allés un peu plus loin et se sont livrés à une analyse marketing de cette réponse. Elodie considérait que *« le service… a été imposé. En gros, la personne n'a pas choisi d'utiliser ce service-là donc ça ne lui a pas procuré de satisfaction ou d'insatisfaction »*, tandis que Nathalie suggérait de *« trouver un attribut pour eux, travailler*

sur un attribut sur lequel ils vont avoir une réaction forte qui va les mobiliser et qui va les faire passer dans la moitié droite, sans quitter cette partie moitié basse ».

Le point « modérément satisfait » / « modérément insatisfait »

Les participants ont effectué deux types d'interprétations de ce point : soit ils l'envisagent comme une évaluation réellement partagée, soit ils l'interprètent comme une difficulté à se prononcer.

Pour Nathalie si *« t'es à la fois "un peu insatisfait" et "un peu satisfait", ton jugement est moyen »*. C'est également ce que pense Lucie qui parlait *« d'opinion mitigée »* ou de Paul qui considérait que *« c'est un séjour "bof bof" dans lequel il n'y avait « pas beaucoup de souvenirs, d'expérience marquante »* (Paul). Louisa a décri plus précisément ce type d'évaluation en ces termes : *« C'était pas transcendantal. Donc, un truc où, par exemple, je suis modérément satisfaite. Il n'y a pas le truc… Ce n'était pas "youpi youpi", ce n'était pas "wahou", et en même temps… Et en même temps, ce n'était pas nul non plus, quoi. Tu vois ? Je n'ai rien de spécifique à reprocher… »*. Elle parle aussi de *« verre à moitié vide à moitié plein »*.

Mais les participants y voyaient aussi *« la volonté de se… De ne pas prendre position »* (Paul), puisqu'il se trouve que *« géographiquement, c'est un refuge »* (Paul). Du coup, pour Nabih *« le point du centre, c'est pour moi, je te l'avais dit tout à l'heure, c'est l'attitude neutre »*, *« c'est quelque chose qui te concerne, mais sur lequel tu n'as pas un avis tranché »*. Pour Anna une telle réponse traduit le fait que *« la personne a répondu de manière totalement neutre en fait. Elle n'a pas voulu se… Elle n'a pas voulu situer sur la question »*. Apportant une analyse plus méthodologique, Paul considérait qu'une personne qui répondrait comme ça était sans doute *« démotivée par l'enquête »*.

Pour conclure, Nabih a repris très clairement les deux interprétations faites concernant ce point dans un verbatim : *« ça peut être dû à plusieurs cheminements mentaux. Tu peux avoir comme je t'ai dit une personne… Moi si je répondais ça, c'est peut-être parce que je m'en fous. C'est pas un truc qui m'intéresse, c'est pas un truc qui m'a marqué, ou alors que je n'arrive pas réellement à dégager de tendance claire. Toutes les expériences que j'ai vécues, satisfaisantes ou pas, s'annulent en quelque sorte. Et que j'exprime ça par ce point-là. Par rapport à celui d'avant par exemple, ça peut être aussi une réflexion sur plusieurs dimensions. Tu as plusieurs dimensions, critères, certains ont été satisfaits, certains ont été pas satisfaits, mais […] t'as pas d'avis tranché sur le point »*.

Le point central de la grille traduirait donc selon les participants une difficulté à se situer clairement sur la question dans la mesure où l'avis n'est pas assez tranché. Autrement dit, ils sont incertains de leur opinion, c'est pour cela qu'ils ne souhaitent pas prendre position.

Le point « extrêmement satisfait » / « extrêmement insatisfait »

Plusieurs participants ont été interpelés par ce point, comme le montre cette remarque d'Elodie : *« C'est marrant parce que c'est une configuration, quand je regardais l'échelle sur la première question, c'est une configuration que je trouvais difficile à appliquer en fait. Ce qui me surprend, c'est que le fait d'être "extrêmement satisfait" par un service pour moi est quasiment incompatible avec une "extrême insatisfaction" »*. La première réaction de Nabih a été de caractériser cette réponse de *« contradictoire ou absurde »* et pour Najate c'était « *très dérangeant »*, *« paradoxal »*.

Une fois cette première impression passée, en essayant de comprendre comment analyser cette réponse, les participants ont tenté de caractériser le type de services auxquels cela pouvait correspondre. Anna a trouvé un exemple dans son expérience personnelle: *« d'un côté le CROUS me satisfait parce que c'est pas cher, c'est pratique et voilà ça, ça me satisfait énormément, mais j'ai beaucoup d'insatisfaction également puisque c'est pas bon, c'est répétitif »*. Pour Elodie *« ça peut s'appliquer à des services qui peuvent être extrêmement utiles à la personne, et peut être qu'elle en espérait beaucoup, et du coup ça provoque… une opinion du coup complexe, composée à la fois de satisfaction et d'insatisfaction »*. Paul est allé un peu plus loin en considérant que : *« c'est les trucs que tu es obligé de prendre, d'utiliser et… Parce que normalement si t'es "extrêmement insatisfait" d'un truc, à moins d'être pomme t'y retournes pas. Donc là ça semble produire un peu le service un peu obligatoire, celui qui ne laisse pas le choix »*.

Ces exemples les ont amenés à associer le choix de ce point de la grille à une forme d'implication vis-à-vis de l'objet évalué. Pour Elodie *« si on va vers "beaucoup", "extrêmement" de satisfaction et "beaucoup", "extrêmement" d'insatisfaction, là je pense que la personne a été très sensible au service de manière générale, qui lui a été rendu, donc c'est quelque chose qu'elle a pris à cœur, et c'est un service qui, peut-être, est suffisamment complexe pour justement générer plein de points de vue »*. Autrement dit, *« la personne est vraiment investie vis-à-vis du… du service fourni »* (Anna), *« ça veut dire que la RATP l'importe »* (Nathalie). Selon Louisa cette implication peut être *« émotionnelle ou ça peut-être rationnel »*.

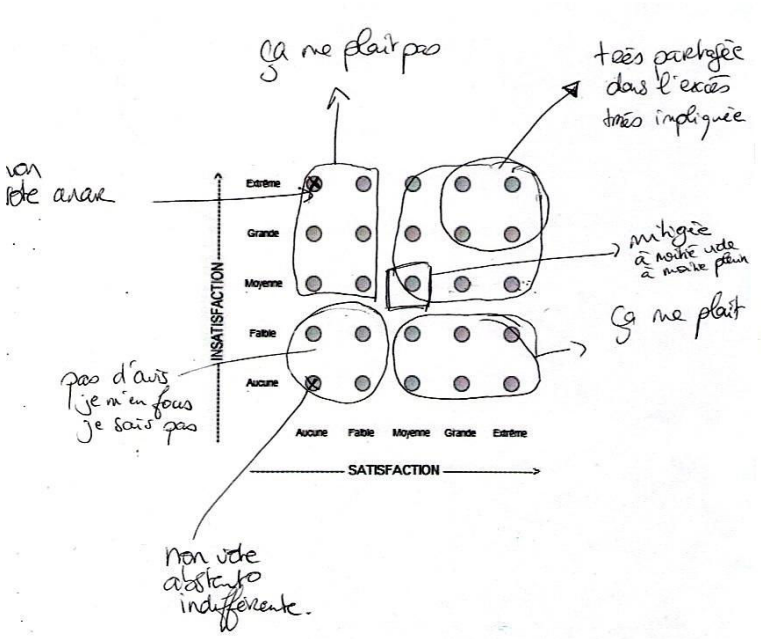
Certains participants ne se sont pas arrêtés au fait qu'il s'agisse d'une forme d'implication. Pour Nabih *« ça c'est assez clairement une zone ambivalente. Tu as une forte satisfaction une forte insatisfaction en même temps »*. Louisa a exprimé les choses ainsi : *« Je dirais que là tu as plein de choses qui rentrent en balance et genre, tu es très ambivalente. Genre tu es… Tu es très partagée je dirais ça. Tu es très partagée parce qu'il y a des trucs qui te réjouissent mais il y a des trucs qui te bloquent totalement. Qui te bloquent… Ouais… Pas vraiment tellement parce que ça te réjouit quand même mais… Mais… Il y a des trucs et de… C'est ça « extrême » / « extrême »… Donc il y a des trucs qui te… Très partagée… Ouais, je suis très partagée et en même temps dans l'excès »*.

Enfin, pour Nathalie une telle réponse présente un potentiel intéressant d'un point de vue marketing : *« les clients sur lesquels il faut, à mon avis, focaliser ses efforts c'est ceux qui sont dans le cadran en haut à droite, puisqu'ils ont un potentiel de satisfaction parce qu'ils sont à droite, mais il y a aussi des critères qui les gênent très fortement, peut être en corrigeant ces quelques critères on arriverait à les faire descendre dans le cadran en bas à droite »*.

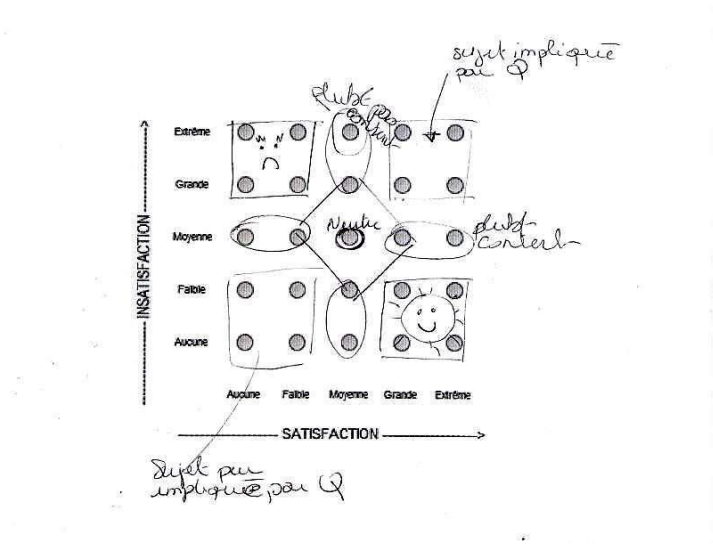
La diagonale dans son ensemble

Pour conclure cette analyse des différentes zones de la grille, il est intéressant de se pencher sur la perception de la diagonale dans son ensemble. A cet égard les participants ont essayé de contraster les points situés en bas de la diagonale et ceux qui se trouvent en haut. Ainsi pour Elodie cette diagonale correspondait à *« une opinion potentiellement… alors "neutre" mais comme on disait tout à l'heure, pas forcément générée de la même manière parce que le "pas du tout" ou "peu"… euh… ce serait plutôt un service qui n'aurait pas procuré d'avis particulier pour la personne, alors que quand on se situe sur le "modérément", "beaucoup" ou "extrêmement", là on est davantage sur une prise de position, de la personne vis-à-vis du service qui lui a parlé davantage en fait. Donc elle se positionne comme étant "satisfaite" mais aussi "insatisfaite" donc ça génère davantage d'opinions »*. Pour Paul *« c'est la diagonale de l'expérience, la diagonale du vécu »*. Il a développé plus précisément son idée dans les termes suivants : *« je dirais que quand on monte, je pense que, enfin, que plus on va vers les bords du carré, euh… C'est entre guillemets de l'expérientiel assez fort. Pas forcément de l'expérience de consommation, mais, un produit impliquant avec une expérience autour quoi. Là tu n'as pas du "pas du tout ; pas du tout" de l'assurance habitation, c'est un truc avec pas d'expérience que tu n'es pas capable d'évaluer. Et puis plus tu vas vers là plus euh… plus il y a du vécu derrière »*.

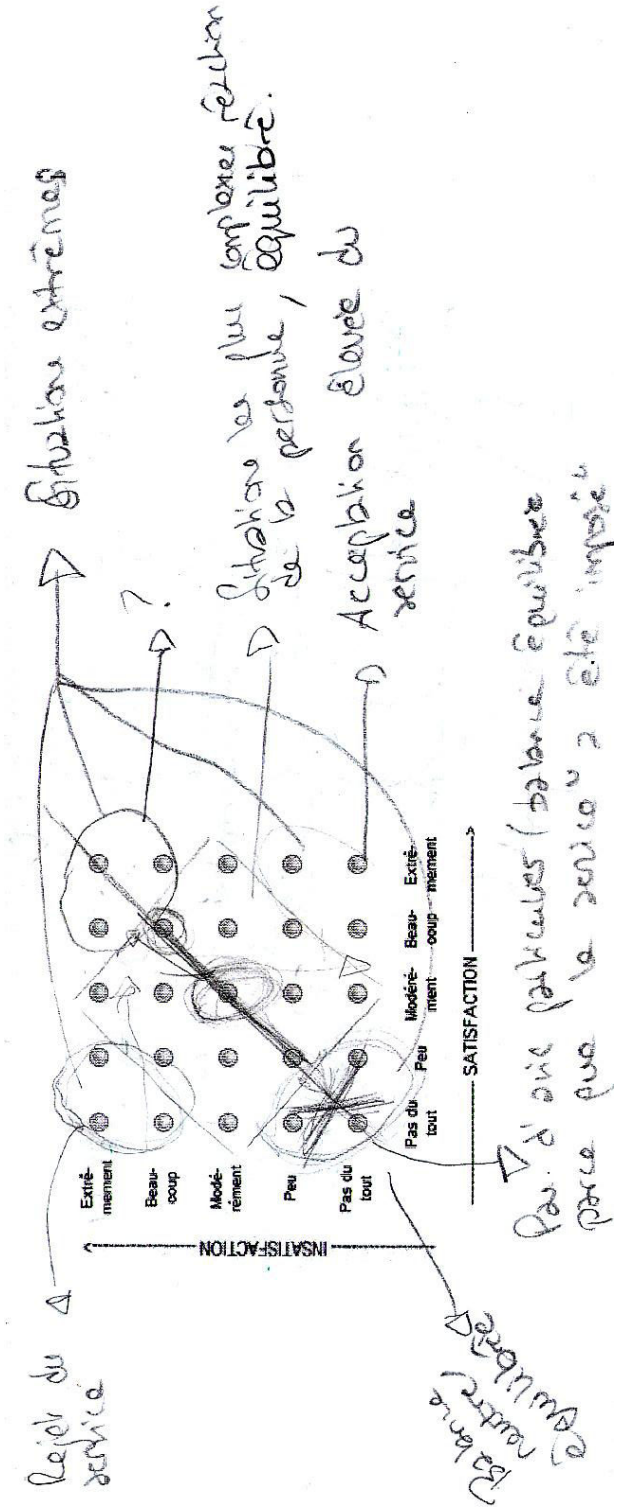
Evaluative Space Grid annotée par Louisa



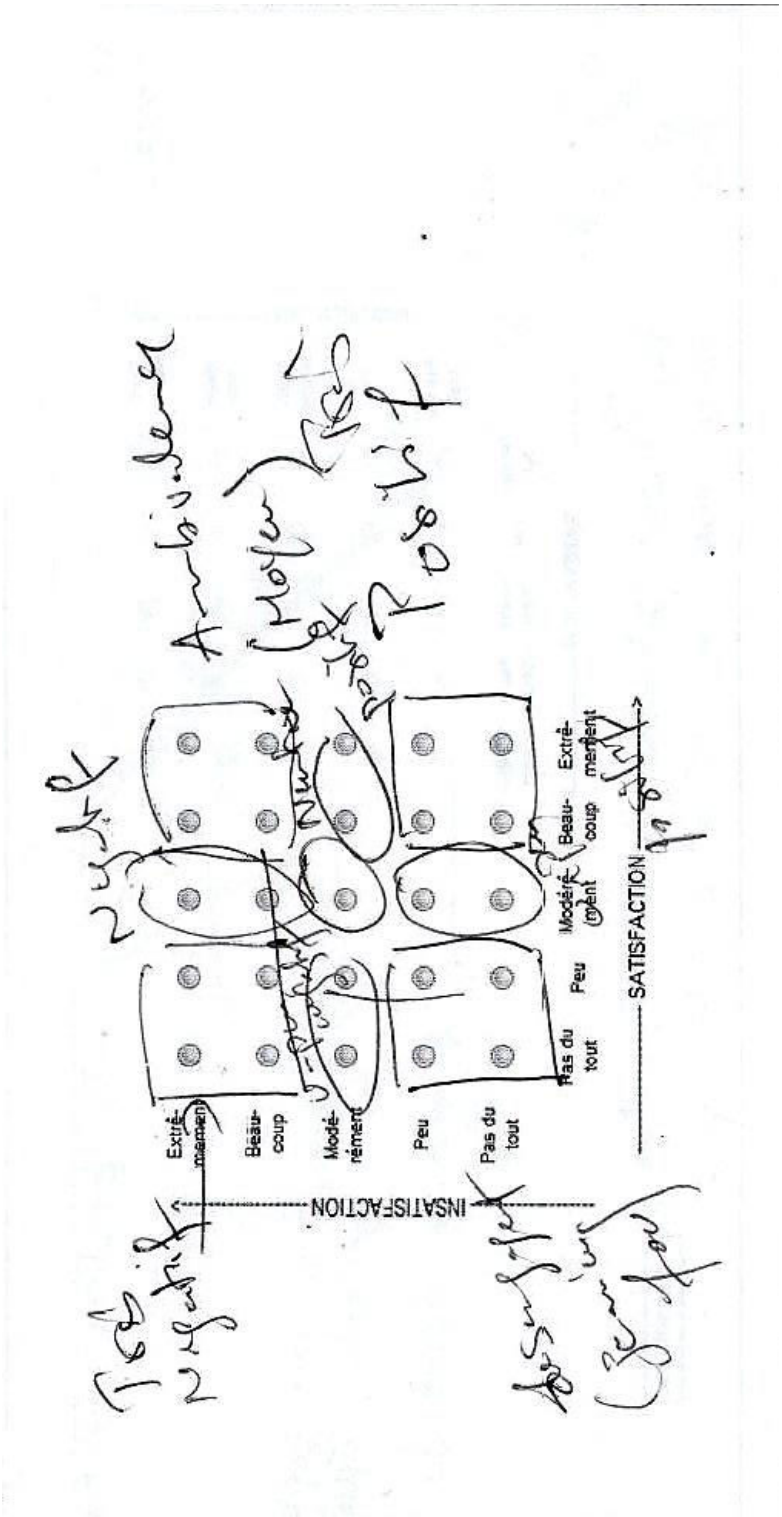
Evaluative Space Grid annotée par Anna



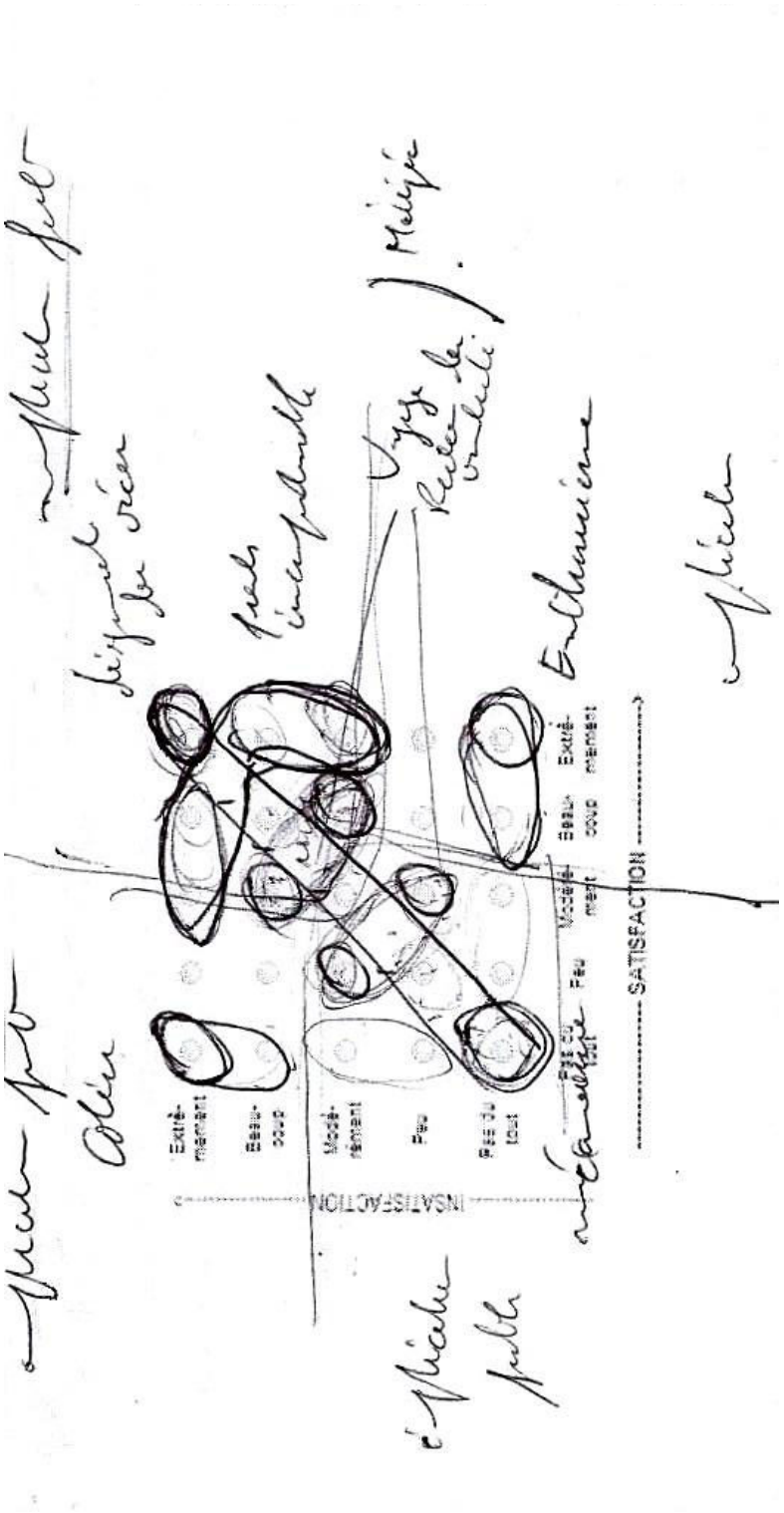
Evaluative Space Grid annotée par Elodie



Evaluative Space Grid annotée par Nabih



Evaluative Space Grid annotée par Paul



Evaluative Space Grid annotée par Najate

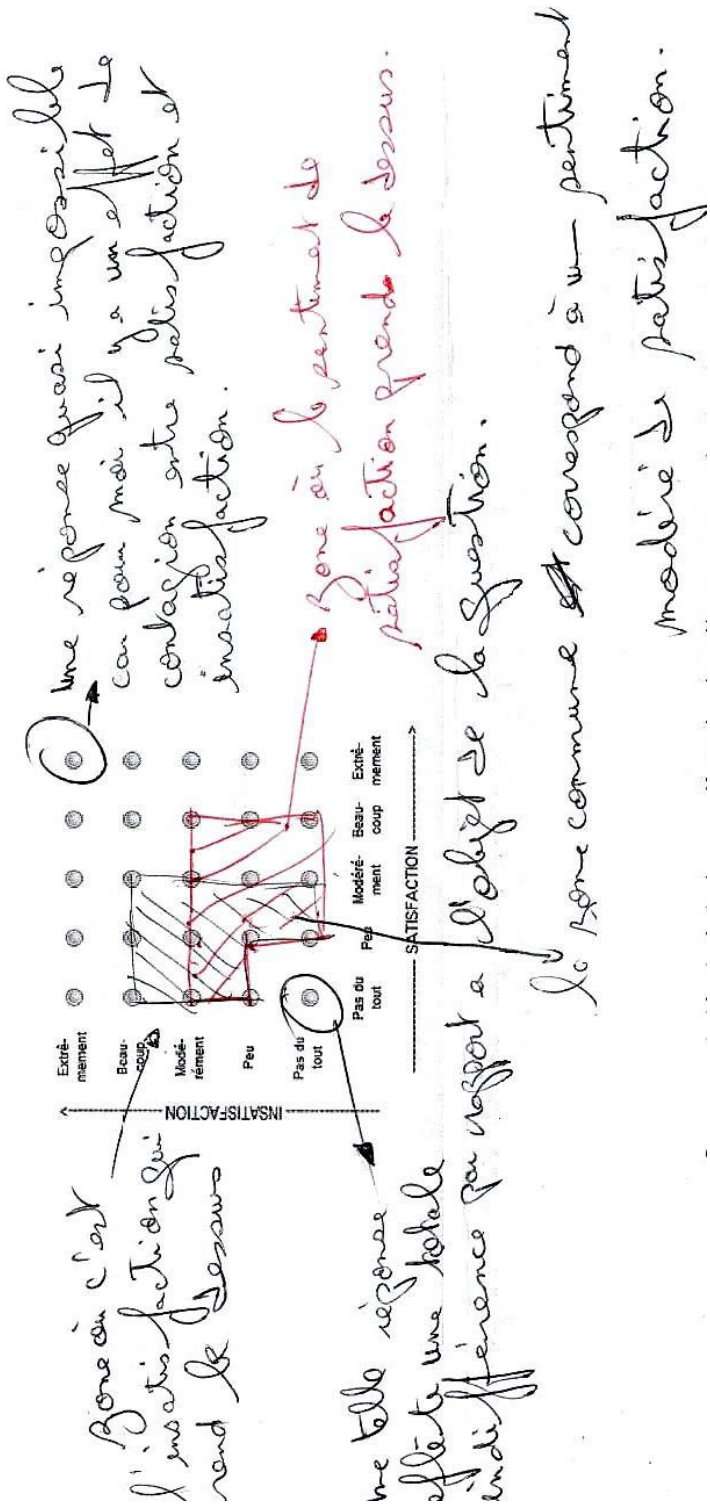


Figure 30. Matérialisation des différentes zones évaluatives de l'outil par les participants

La figure 30 présente des grilles Larsen annotées par certains des participants pour définir les différentes zones évaluatives de l'outil.

Point « pas du tout satisfait » ; « pas du tout insatisfait »	Point « modérément satisfait » ; « modérément insatisfait »	Point « extrêmement satisfait » ; « extrêmement insatisfait »
« *ne pense rien* » « *pas d'expérience* » « *pas d'avis* » « *pas d'émotion* » « *pas impliqué* »	« *opinion mitigée* » « *arrive pas à dégager une tendance claire* » « *pas d'avis tranché* » « *pas voulu se situer sur la question* » « *refuge* »	« *très partagé et même temps dans l'excès* » « *suffisamment complexe pour générer plein de point de vue* » « *quelque chose pris à cœur* » « *ça importe* »
Diagonale dans son ensemble		
« *diagonale de la neutralité* ; « *diagonale de l'expérience* » ; « *diagonale du vécu* » Pour décrire le fait de monter le long de la diagonale : « *prise de position* » ; « *[vers de] l'expérientiel assez fort* » ; « *[vers de] l'implication* »		

Tableau 31. Résumé des interprétations associées aux différents points de la diagonale

Le tableau 30 synthétise le sens attribué aux différents points de la diagonale de la grille de Larsen dans le discours des participants.

3.4. Intérêt perçu de l'*Evaluative Space Grid*

Unanimement, les participants ont reconnu l'intérêt de *l'Evaluative Space Grid*. Plusieurs éléments sont ressortis de leurs discours. Tout d'abord, l'outil permet selon eux une visualisation spatiale des évaluations, facilitant l'interprétation et la communication des appréciations. Ensuite, ils ont envisagé la grille comme un outil permettant d'obtenir des informations plus riches et plus fines qu'un différentiel sémantique classique. Ils pensent aussi que cet outil alternatif aide à structurer les réponses obtenues et peut être utilisé comme un support pour approfondir une évaluation. Enfin ils nous ont suggéré plusieurs pistes d'utilisation.

Un outil permettant de « spatialiser » les évaluations

Le caractère géométrique de l'outil a inspiré les participants. Pour Nathalie *« ça a l'avantage […] de proposer une visibilité graphique, plus facile à comprendre que des échelles linéaires je*

trouve ». Paul pense que *« la représentation très géométrique de la chose »* suscite *« des connotations spatiales. Tu vois le fait que tu aies des coins, quand je disais tout à l'heure "moi je ne m'enflamme pas", je m'éloigne des coins quoi, tu vois je ne prends pas de position extrême… »*. Il nous a même suggéré *« à partir des réponses linéaires que tu as eues tu peux calculer les distances entre les points ou faire une représentation anamorphosée*[11] *»*.

Un outil permettant de gagner en information

Tous les participants s'accordaient sur un point : la grille permet de *« générer davantage d'informations précises sur le ressenti des gens »* (Elodie), *« ça permet de donner un jugement un peu plus fin qu'un truc linéaire »* (Nabih). Nathalie a illustré plus précisément comment cela pourrait être concrètement utilisé en marketing : *« par exemple sur les parfums, ça serait vachement intéressant de retravailler les parfums avec ça parce que ça permet de… de pointer du doigt qu'il peut y avoir des choses opposées qui se passent et qui sont masquées dans un jugement global […]. Si ça se trouve sur leur parfum, s'ils passaient juste à une mesure comme ça, c'est pas très compliqué, s'ils passaient juste à une mesure comme ça, ils prédiraient mieux les intentions portées »*.

Un outil permettant de structurer l'évaluation pour ouvrir la discussion

La grille semblait aider certains répondants à structurer leur réponse en réfléchissant aux aspects positifs et négatifs de leur appréciation. C'est ce dont témoignait Lucie dans le verbatim suivant : *« même si c'est un peu difficile, ça m'amène à mieux structurer ce que je ressens »*. Pour Elodie *« ça oblige à se poser plein de questions. Beaucoup plus que… La question classique êtes-vous plutôt satisfait ou plutôt insatisfait, parce que avec la question sur le continuum, on donne un avis global alors que là on donne peut-être un avis global, mais on va se poser la question qu'est-ce qui m'a procuré de la satisfaction et qu'est ce qui m'a procuré de l'insatisfaction ? »*.

Certains ont également envisagé la grille comme *« une amorce pour faire parler […], pour dire mais pourquoi êtes-vous "très insatisfait" pourquoi êtes-vous "très satisfait" ? »* (Nathalie). On retrouve cette même idée dans les propos d'Elodie qui considérait que *« c'est un outil qui doit être complété, par justement des questions qualitatives derrière en disant : "voilà, vous êtes positionnés sur cette matrice, qu'est-ce qui explique votre niveau de*

[11] *Une anamorphose est une déformation réversible d'une image dont l'objectif est d´adapter la forme de l'objet représenté non pas à la réalité physique mais plutôt à la réalité perçue. Cette déformation de l´espace peut être construite par un modèle mathématique à partir de données quantitatives.*

satisfaction et qu'est-ce qui explique notre niveau d'insatisfaction ?" Et pour moi [...] ça peut permettre d'avoir un outil super complet, beaucoup plus précis en termes de mesure du ressenti des gens vis-à-vis de leur expérience ».

Un outil offrant différentes voies d'applications

Enfin, les participants nous ont suggéré des pistes d'utilisation de la grille de Larsen. Nabih a proposé de ne pas la limiter la désignation des axes à l'évaluation d'un seul concept, mais de les utiliser pour comparer l'importance relative de plusieurs critères : *« Un autre truc, je pense que ça peut être sympa de l'utiliser sur des items. [...] Je pense au fait d'aller au restaurant. Je vais comparer le service avec la bouffe, avec le prix, avec la décoration [...]. Je peux être quelqu'un qui s'en fout du décor. Et le fait que le décor je trouve ça laid, n'impacte pas beaucoup, alors que s'il y avait quatre questions… Comment trouvez-vous le décor, la bouffe etc. ».* Son idée était de mettre en balance des critères les uns par rapport aux autres, en les combinant dans le cadre de la grille plutôt que d'effectuer des évaluations critère par critère de manière à isoler les attributs qui ne sont pas pertinents pour certains répondants.

Louisa, quant à elle, a eu *« vraiment très envie d'appliquer ce truc politiquement ».* Elle a exprimé son idée en ces termes : *« Voilà abstention anarchiste et abstention indifférente. Mais je ne sais pas pourquoi je pensais à ça tout à l'heure, parce qu'en fait, je me dis que j'aurais pu potentiellement… Aussi tu vois… Par exemple pour le PS je pense que si tu me demandais juste au moment où ils sont passés, et aujourd'hui, par exemple, et bien j'aurais pu évoluer dans ma grille, tu vois ? J'aurais pu me déplacer. Je pense que tu vois par exemple, si… Je pense que si j'avais été… Je ne sais pas, j'essaie de faire l'exercice ? Genre, juste le jour de l'élection, tu vois, je pourrais dire que la satisfaction était grande, et l'insatisfaction moyenne, j'aurais été là ou là ».* Elle nous a ensuite montré que son évaluation s'était déplacée vers la droite de la grille, en raison de la baisse de son niveau de satisfaction à l'égard du gouvernement en place.

Il y avait en fait, dans son discours, deux idées distinctes. D'une part, elle considérait que la grille de Larsen pouvait être intéressante pour analyser plusieurs comportements identiques qui résultent d'opinions différentes. En effet, le fait de ne pas voter pour un anarchiste s'explique par une évaluation très négative du système politique, alors que d'autres formes d'abstention peuvent être liées à une absence d'opinion ou un mélange d'opinions positives et négatives à l'égard des candidats. Les questions d'indifférence et d'ambivalence

sont d'ailleurs étudiées en sciences politiques (Thornton, 2011 ; Yoo, 2010). D'autre part, elle pensait que la grille pourrait être utilisée pour étudier des parcours d'évolution des opinions.

4. Conclusion

Le regard de ces chercheurs apporte des clés de compréhension concernant la signification attribuée aux différentes zones de la grille de Larsen. Ces éléments permettent, par triangulation des méthodologies, de confirmer les conclusions issues de la phase quantitative (Denzin, 1978). Ils font également émerger des nouvelles propositions.

Tout d'abord, nos craintes concernant d'éventuelles difficultés d'utilisation de la grille de Larsen sont confirmées. Cela nous permet de valider l'hypothèse, formulée dans la discussion de l'étude 3 (Chapitre 4), d'un taux d'erreurs plus élevé avec l'*Evaluative Space Grid* qu'avec les différentiels sémantiques. Toutefois, ce constat ne doit pas se traduire par le rejet de l'outil. De multiples pistes d'amélioration de la présentation de la grille de Larsen ressortent également du discours des répondants. Par exemple, indiquer le long des axes qu'il s'agit d'évaluer les « éléments » ou « attributs » perçus positivement (resp. négativement) rendrait la compréhension de l'outil plus immédiate en le positionnant directement comme un outil d'évaluation globale et multicritères.

Cette première partie de l'étude est également riche d'enseignements concernant la manière dont les individus parviennent à former une évaluation globale lorsqu'ils utilisent la grille. Celle-ci résulte « mécaniquement » de la combinaison entre les notes attribuées sur chaque axe. En ce qui concerne le passage d'un point à un autre sur les axes, il semble pouvoir résulter aussi bien du nombre d'attributs pris en compte dans l'évaluation que de leur importance. Cette indistinction indique clairement que c'est au niveau de la note donnée sur chaque axe que l'évaluation est globalisée.

En ce qui concerne l'interprétation faite par les répondants des différents points de la grille de Larsen, l'apport de ces entretiens semi-directifs repose principalement sur l'analyse de la diagonale. Voici les principaux éléments qui se dégagent :

- La zone basse de la diagonale (située en bas à gauche de la grille de Larsen) est associée à une absence d'opinion liée à un manque de réaction vis-à-vis de l'objet évalué ;

- Le point central de la grille, qualifié de « refuge » par l'un des participants, est plutôt associé à la volonté de ne pas se prononcer sur la question posée, soit parce que l'opinion est réellement partagée, soit parce qu'elle n'est pas assez claire dans l'esprit du répondant. Il semble que ce point soit associé à une forme de doute ou d'incertitude quant à la réponse à donner ;

- Quant à la zone haute de la diagonale (située en haut à droite de la grille de Larsen), elle est associée à des réactions ambivalentes de conflit intense, mises en lien avec de « *l'implication* », du « *vécu* » et de « *l'émotionnel* » à l'égard de l'objet évalué.

Ensuite, cette première partie de l'étude qualitative permet de mettre en évidence plusieurs pistes d'utilisation de la grille. L'un des participants nous a notamment suggéré de capitaliser sur le fait que l'outil favorise la structuration de l'évaluation pour l'utiliser comme un support à la discussion pour mieux comprendre quels sont les éléments qui suscitent de la satisfaction et de l'insatisfaction.

Pour délimiter plus clairement les contextes dans lesquels la grille pourrait constituer un outil de mesure pertinent, nous avons complété ses entretiens de chercheurs par des interviews de professionnels familiers avec les études et les avis de consommateurs en ligne.

Section 2. Qu'en pensent les professionnels ?

De manière à ancrer davantage notre travail dans la réalité de terrain, nous avons soumis *l'Evaluative Space Grid* à l'expertise de cinq professionnels de l'industrie des études et des avis de consommateurs en ligne.

1. Présentation de l'étude et méthode

1.1. Nombre d'entretiens et choix de l'échantillon

Nous avons interrogé cinq personnes confrontées dans leur métier au problème de la mesure des évaluations de consommateurs. Deux profils de personnes ont été sollicités : des professionnels des études et des spécialistes des avis de consommateurs en ligne. Le tableau 31 ci-dessous décrit plus précisément les fonctions occupées par les personnes

interrogées. Sur demande de certains participants, leur nom a été modifié. En accord avec eux, leurs propos sont donc rapportés sous un prénom d'emprunt.

Participants	Fonction et entreprise
Julie	Chargée de projets chez TNS Sofres, secteur tourisme, culture, loisirs ; ancienne chargée d'études sénior
Helen Zeitoun	Responsable mondiale « Brand and Customer Experience » chez GfK ; ancienne Présidente de l'Association Française du Marketing
Pierre Gomy	Directeur Marketing chez Millard Brown, société d'études spécialisée dans l'analyse de l'image de marque
Séverine Philardeau	Vice-présidente « Global Partnerships » chez TripAdvisor
Gilles Granger	Fondateur et directeur général de Vinivi, société fondée en 2006 et spécialisée dans les avis vérifiés de voyageurs

Tableau 32. Présentation des participants, de leur fonction et de leur entreprise

1.2. Déroulement des entretiens

Tous les entretiens se sont déroulés par téléphone, sauf celui avec Julie, réalisé en face à face. Ils ont duré entre une demi-heure et une heure. En guise d'introduction, les entretiens commençaient par quelques questions sur le parcours professionnel des participants. Nous abordions ensuite les thèmes suivants :

- Les pratiques de mesure des évaluations globales dans le cadre de leur travail ;
- La question de la prise en compte de l'ambivalence des répondants ;
- Leur réaction face à l'*Evaluative Space Grid* : intérêt et prise en main ;
- Les perspectives potentielles d'application.

Pour les entretiens réalisés par téléphone, un lien vers un questionnaire utilisant l'outil a été envoyé durant la conversation de manière à recueillir la réaction spontanée des participants. Le guide d'entretien est présenté en annexe 5.

2. Analyses et résultats

Chaque entretien a fait l'objet d'un compte rendu détaillé. Compte tenu de leurs fonctions, les participants disposaient d'un temps limité. Dans le but d'optimiser le temps qui nous était accordé, nous avons adopté une technique d'interview relativement directive. En conséquence, les thèmes retenus lors de l'analyse sont ceux du guide d'entretien.

2.1. La mesure des évaluations globales

Les outils de mesure utilisés

Dans les études de satisfaction ou de perception, l'utilisation d'une mesure globale est systématique. TNS Sofres propose par exemple à ses clients un indicateur de performance globale incluant une mesure de satisfaction. Cet indicateur permet d'attribuer un score à la marque qui peut ensuite servir de benchmark pour effectuer des comparaisons avec les marques concurrentes. Toutefois, lorsque les sociétés d'études construisent un questionnaire, elles doivent s'adapter aux objectifs et aux désirs du donneur d'ordre. Il y a donc autant d'échelles qu'il y a de clients. Malgré tout, chez TNS Sofres, la satisfaction globale est généralement mesurée avec une échelle en 5 ou 10 points allant « pas du tout satisfait » à « tout à fait satisfait ».

Lorsqu'il s'agit de recueillir des évaluations de consommateurs en ligne, les formats sont également très divers (cf. Chapitre 1, Section 3). L'échelle utilisée par Vinivi est composée de 5 points allant de « très insuffisant » à « très satisfaisant ». Ces points sont accompagnés d'un code couleur allant du rouge au vert. De même, TripAdvisor utilise des échelles composée de 5 points verts allant de « horrible » à « excellent » (figure 31).

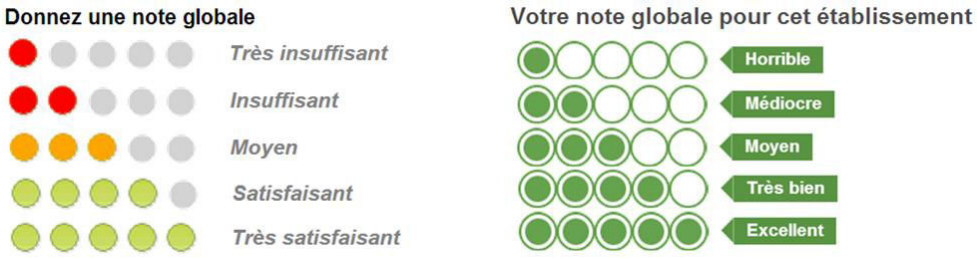


Figure 31. Dispositif de mesure utilisé par Vinivi (à gauche) et par TripAdvisor (à droite)

Gilles Granger ainsi que Séverine Philardeau ont justifié le choix d'une échelle en 5 points par la volonté de privilégier des outils simples. En effet, pour Gilles Granger les échelles en 10 points augmenteraient à tort la variance interindividuelle : certains individus attribueraient une note de 7 lorsqu'ils sont très satisfaits alors que d'autres choisiraient 9. Ils justifient également le fait d'associer une étiquette sémantique à tous les points de l'échelle. Gilles Granger considère que les mots sont plus parlants que les chiffes et pour Séverine Philardeau cela permet de « cadrer l'évaluation ». Cette préoccupation est importante pour TripAdvisor dans la mesure où il s'agit d'un site international. Dans cette perspective, les chiffres ont pour objectif de réduire les différences culturelles. Séverine Philardeau a toutefois insisté sur le fait que ce dispositif de mesure est à envisager comme un outil subjectif. Par exemple, un commentaire textuel moyen peut être associé à une note de 2/5 comme à une note de 4/5. Cette « subjectivité » des notes est lissée par le nombre élevé d'évaluations.

La qualification du point central

Les professionnels des études connaissent bien la problématique du point central des échelles. L'étiquette sémantique associée à ce point est variable, selon les objectifs de l'étude et les exigences du client. Chez TNS Sofres, les points intermédiaires ne sont pas toujours qualifiés. Lorsqu'ils le sont, c'est généralement l'étiquette sémantique « ni satisfait ; ni pas satisfait » qui est utilisée, plus rarement il peut s'agir du terme « moyennement ». Julie a également évoqué le qualificatif « probablement » dans le cadre de la mesure de l'intention de réachat. Pierre Gomy, quant à lui, nous a parlé de point « neutre ». Chez TNS Sofres comme chez Millward Brown, une modalité « ne sais pas » est parfois prévue séparément. En ce qui concerne les sites de recueil d'avis de consommateurs, Vinini comme TripAdvisor utilisent l'étiquette sémantique « moyen », comme l'indique la figure 31. Toutefois chez Vinivi, le point central est de couleur orangée ce qui, selon Gilles Granger, lui associe une connotation négative ».

L'analyse du point central

La plupart du temps, les réponses associées au point central sont ignorées de l'analyse. Pour illustration, Julie a évoqué le souvenir d'une étude dans le secteur du tourisme dans laquelle une part importante de répondants avait coché le point central (le « 5 » sur une échelle allant de « 1 » à « 10 »). Ne sachant pas comment les analyser (répondants désimpliqués par l'enquête ? méconnaissance du service évalué ? évaluation moyenne ?),

les responsables de l'étude avaient décidé de ne pas exploiter ces réponses. Cette pratique, très courante, s'apparente aux préconisations de Reichheld (2003) qui propose de calculer le *Net Promoter Score* en retranchant la part de « détracteurs » (les répondants insatisfaits) aux « promoteurs » (les répondants satisfaits). Dans cette perspective, les répondants ayant choisi le point central, qualifiés de « passifs » par Reichheld (2003) sont ignorés de l'analyse (cf. Chapitre 2, Section 2). Gilles Granger nous a précisément expliqué qu'il utilise cette technique pour « sortir de l'ambivalence ».

2.2. Prise en compte de l'ambivalence

Helen Zeitoun a insisté sur l'importance de la question de l'ambivalence et sur la nécessité de la prendre en compte dans les études consommateurs. Selon elle, l'expérience de consommation et le rapport à la marque sont des domaines dans lesquels les réactions ambivalentes sont très présentes. La journée du consommateur est pleine d'hésitations et de choix à effectuer sous la pression de l'ambivalence : acheter responsable mais pas cher, savoureux mais pauvre en calories… Pour elle, le marketing manager est un gestionnaire d'ambivalence.

Les solutions mises en œuvre pour saisir l'ambivalence

Les professionnels interrogés nous ont unanimement dit qu'ils ne prévoyaient pas de mesure spécifique de l'ambivalence dans leurs questionnaires. Toutefois, ils ont tous considéré qu'elle pouvait être prise en compte grâce à la mesure détaillée des attributs qui suit généralement la note globale (cf. figure 32, illustrant les items utilisés à cette fin par Vinivi). Par exemple, lorsque Millward Brown réalise une étude sur un territoire de marque, les items utilisés se regroupent pour former plusieurs dimensions. Certains facteurs peuvent alors être évalués positivement et d'autres négativement. Cela peut être envisagé comme un moyen indirect de prendre en compte d'éventuelles réactions ambivalentes.

Pour Séverine Philardeau, s'il est vrai que la note globale, telle qu'elle est mesurée par TripAdvisor ne permet pas de saisir l'ambivalence, l'analyse du commentaire textuel qui l'accompagne peut être envisagée comme un moyen de résoudre ce problème. Pour ce faire, il existe des d'outils d'analyse sémantique permettant de décrypter le sens du commentaire.

Gilles Granger a, quant à lui, proposé de « se débarrasser » de l'ambivalence en utilisant la mesure binaire de l'intention de recommandation. Cette question permet de classer les

individus en deux groupes : ceux qui recommandent et ceux qui ne recommandent pas, évacuant ainsi le problème de la mesure de l'ambivalence.

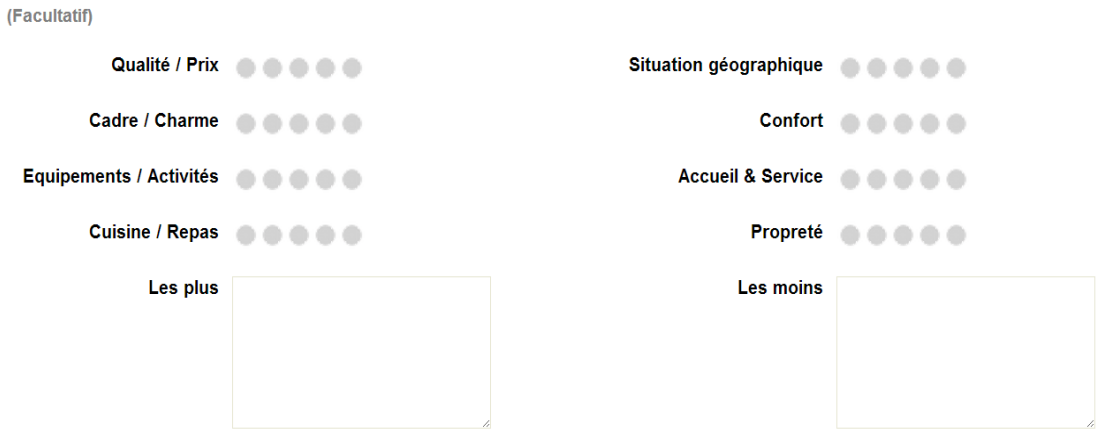


Figure 32. Items utilisés par Vinivi pour l'évaluation d'un hôtel

2.3. Réactions face à l'*Evaluative Space Grid*

Globalement les réactions des professionnels des études ont été beaucoup plus favorables que celle des spécialistes des avis de consommateurs en ligne.

Le point de vue des professionnels des études

La première réaction de Julie a été de nous dire qu'elle était « très convaincue » par l'outil. Elle considérait que l'*Evaluative Space Grid* permettrait d'augmenter les nuances et de collecter une information plus précise. Pierre Gomy l'a immédiatement opposée au *Net Promoter Score*, technique qu'il juge « très pauvre » dans la mesure où elle se limite à une analyse des notes extrêmes. Selon lui, la grille de Larsen est à envisager comme un moyen d'inciter les répondants à développer une note globale initialement donnée sur un différentiel sémantique ou une échelle de Likert. Enfin, Helen Zeitoun a trouvé l'*Evaluative Space Grid* « très intéressante », abondant dans le sens de Pierre Gomy concernant la pauvreté des outils existants pour saisir les évaluations globales. Elle a vu dans la grille de Larsen un moyen de « sortir des échelles pour aller vers des outils plus graphiques ». Dans cette perspective, elle a proposé de présenter l'outil sous forme de cadrans mis en évidence par des couleurs pour distinguer les zones d'évaluation positive, négative, indifférente et ambivalente.

Les professionnels des études n'étaient pas inquiets quant à l'utilisation de l'*Evaluative Space Grid* par les panélistes. Julie nous a expliqué que chez TNS Sofres les enquêtes réalisées peuvent être longues, jusqu'à 40 minutes. Cela rendrait possible l'ajout de quelques explications pour introduire un nouvel outil. Dans cette perspective, il lui semblerait important d'expliquer que l'évaluation globale se construit à partir de l'ensemble de l'expérience (confort, service…) et de donner un exemple d'évaluation ambivalente effectuée avec la grille.

Pour Julie, l'*Evaluative Space Grid* est destinée aux enquêtes en ligne ou à toutes les études assistées d'un enquêteur. Toutefois, elle a considéré que l'outil ne serait pas utilisable au téléphone compte tenu de la nécessité de visualiser l'échelle pour pourvoir répondre. Or, les enquêtes CATI (*computer-assisted telephone interview*) constituent encore une part non négligeable de l'activité, surtout pour les études de satisfaction.

L'outil a toutefois soulevé quelques questions en termes d'analyse des données collectées. Julie a cherché à savoir si la grille permettrait de réaliser les tests statistiques habituellement effectués. En d'autres termes, elle a soulevé la question de son codage. Nous lui avons alors présenté les options mises en œuvre dans la thèse (extraction des coordonnées de satisfaction et d'insatisfaction réalisée dans les études 1 et 2 et codage bilinéaire de l'étude 3), ce qui a semblé la rassurer quant aux possibilités d'analyses réalisables avec cet outil. Pierre Gomy s'est plus précisément interrogé sur le pouvoir prédictif de *l'Evaluative Space Grid*. Il y a selon lui une certaine résistance au changement dans l'industrie des études. Pour convaincre le secteur d'adopter un nouvel outil, il faut être en mesure de démontrer qu'il améliore le pouvoir explicatif des variables marketing d'intérêt.

Le point de vue des spécialistes des avis de consommateurs en ligne

Séverine Philardeau et Gilles Granger n'ont pas partagé l'enthousiasme des professionnels des études. La première a jugé l'outil trop compliqué et considéré qu'il ne correspondait pas aux objectifs de TripAdvisor. En effet, les sites d'avis de consommateurs ont besoin d'outils simples et rapides permettant de synthétiser d'importants volumes de données. Par ailleurs, les avis de consommateurs étant déposés volontairement sur ce type de plateforme, les évaluateurs indifférents sont rares. Cela limite considérablement la plus-value d'une distinction entre indifférence et ambivalence selon Séverine Philardeau.

D'autre part, les deux spécialistes des avis de consommateurs en ligne se sont montrés sceptiques concernant le succès de la prise en main d'un tel outil par les internautes. Pour Séverine Philardeau, c'est avant tout le nombre de points qui risquerait de dérouter les répondants. Pour Gilles Granger, l'outil risquerait d'augmenter fortement le taux d'abandon des évaluateurs. Il faut souligner que cette préoccupation a été abordée par tous les professionnels interrogés.

Séverine Philardeau et Gilles Granger envisageaient plutôt l'*Evaluative Space Grid* comme une évaluation qui viendrait en support à un entretien qualitatif. Pour eux, cet outil ne peut s'envisager que si l'on a du temps. Dans ce cadre, la note issue de la grille pourrait être utilisée comme un indicateur chiffré permettant de synthétiser les résultats d'une étude qualitative.

2.4. Domaines d'application envisagés

La plupart des participants ont immédiatement pensé à une utilisation de l'outil pour mesurer la satisfaction client à l'égard des services. Julie et Pierre Gomy ont évoqué la téléphonie mobile, secteur dans lequel l'évaluation globale repose sur des attributs très distincts et donc propice à la formation d'évaluations ambivalentes. Pour Séverine Philardeau, l'outil pourrait être intéressant pour évaluer des services de e-commerce. En effet, dans ce domaine les différents attributs d'évaluation peuvent être complètement déconnectés les uns des autres : on peut être très satisfait du service client et de la livraison et totalement déçu par le produit commandé. Globalement, elle considérait que l'*Evaluative Space Grid* pourrait être utile dans toutes les situations dans lesquelles la dimension produit et la dimension service sont très déconnectées. Par exemple en Angleterre il y a plusieurs fournisseurs de gaz et d'électricité. On peut être satisfait du produit brut acheté mais pas du tout du service (facture incompréhensible, hotline injoignable…). Dans ces cas-là il peut être intéressant d'avoir plus de détails sur le niveau de satisfaction et le niveau d'insatisfaction.

Pierre Gomy a envisagé une utilisation de l'outil dans une perspective d'évaluation de l'image de marque. Selon lui, il y a par exemple beaucoup d'ambivalence vis-à-vis des marques de voitures. Il nous a donné l'exemple de BMW qui capitalise sur un mélange d'évasion et de transgression. Certains consommateurs sont tiraillés entre l'attrait qu'ils ont pour les performances techniques de ces voitures et l'image sociale que renvoie la marque. C'est une marque « voyou », selon les termes de Pierre Gomy. Pour garder sa clientèle, elle

doit maintenir le juste équilibre entre ces deux aspects. Dans cette perspective, l'*Evaluative Space Grid* pourrait être intéressante pour faire des tests de publicité. Cela serait un moyen de vérifier que la publicité a effectivement déclenché l'ambivalence recherchée par l'annonceur. En effet, certaines marques communiquent directement sur l'ambivalence. A cet égard, Pierre Gomy nous a donné l'exemple d'une publicité dans laquelle un jeune marié devait arbitrer entre un monospace et un coupé-sport. L'annonceur avait choisi de mettre directement en scène l'ambivalence de ces clients potentiels en proposant de la résoudre en choisissant un coupé-sport offrant une bonne habitabilité. L'utilisation de l'*Evaluative Space Grid* dans le cadre de tests publicitaires a aussi été suggérée par Helene Zeitoun.

Enfin, Pierre Gomy nous a également cité d'autres domaines qu'il estime être par nature porteur de conflits et pour lesquels il pourrait s'avérer intéressant de saisir la part d'ambivalence des consommateurs : le passage d'un cycle de vie à un autre, la question du maquillage pour les hommes ou encore les assurances qui sont en lien avec les accidents ou la mort.

Section 3. Conclusion

Cette étude qualitative permet d'une part, de clarifier le sens attribué aux différentes zones évaluatives de la grille de Larsen et d'autre part, de limiter l'outil à une utilisation dans le cadre d'études consommateurs.

En ce qui concerne le sens donné aux différents points de la grille de Larsen, l'analyse est relativement conforme, d'une part aux travaux de Larsen et de ses collègues (2009) et d'autre part à la lecture que nous en avons faite jusqu'ici. Elle souligne néanmoins l'importance du point central de la grille, qui, contrairement à ce que nous avons envisagé jusqu'ici, ne relève pas uniquement de l'ambivalence, mais avant tout du « refuge » ou d'évaluations pour lesquelles les répondants n'arrivent pas « à dégager une tendance claire ». Nous proposons donc de distinguer maintenant cinq zones évaluatives : une zone de satisfaction, une zone d'insatisfaction, une zone d'indifférence, une zone d'ambivalence et une zone d'incertitude.

Plus précisément, nous avançons que la diagonale de l'*Evaluative Space Grid* va de l'absence de réaction (ou indifférence) à la cooccurrence de réactions positives et négatives fortes (ambivalence), en passant par la volonté de ne pas trancher. En effet, le point central

de la grille a été associé par les participants à des réactions qui ne sont pas suffisamment fortes positivement ou négativement pour que le répondant décide de se déplacer vers les autres cadrans de la grille. Nous émettons l'hypothèse que ce point attire les réponses incertaines.

La figure 33 représente l'interprétation spatiale que nous proposons des différentes zones évaluatives de la grille de Larsen.

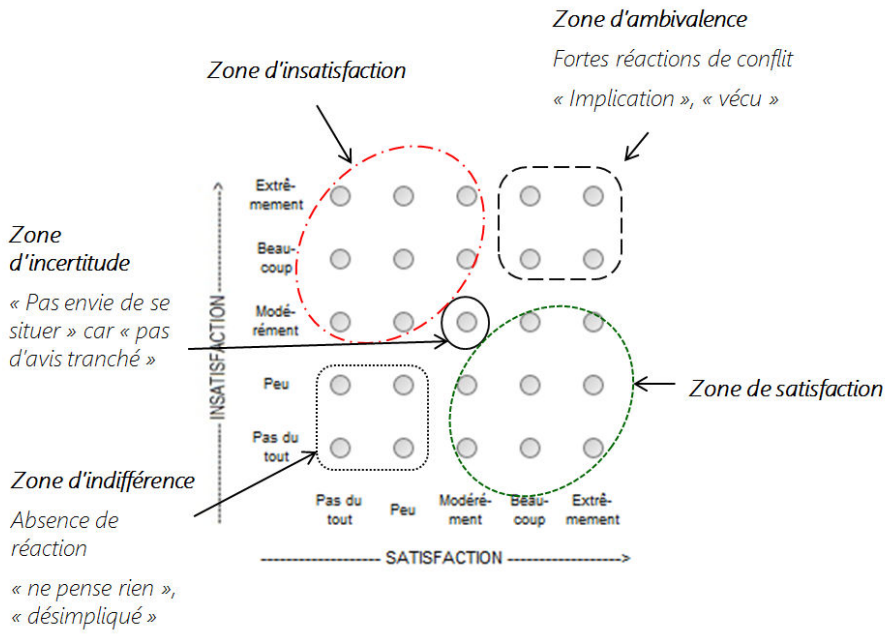


Figure 33. Proposition d'interprétation des différentes zones évaluatives de la grille de Larsen

En ce qui concerne les apports des entretiens réalisés avec les professionnels, ils nous ont permis de limiter clairement l'application de l'*Evaluative Space Grid* au domaine des études consommateurs. De cette perspective, l'outil pourrait être envisagé dans des formats longs, qu'il s'agisse de questionnaires ou d'entretiens. La grille de Larsen pourrait alors servir à développer une note donnée sur un outil plus condensé ou à étayer des arguments dans le cadre d'une interview. En ce qui concerne les domaines d'application, la question de la mesure de l'ambivalence à l'égard des marques semble être une piste intéressante à creuser.

***

Le Chapitre 6 de la thèse vise à tester les propositions issues de cette phase qualitative dans le cadre d'une enquête de satisfaction (étude 5).

Les gens ne cessent de dire qu'il est beau d'avoir des certitudes. Il semble qu'ils aient complètement oublié la beauté bien plus subtile du doute. Croire est tellement médiocre. Douter est tellement absorbant. Rester vigilant, c'est vivre ; être bercé par la certitude, c'est mourir.

Oscar Wilde, « Aphorismes », Editions Mille et une nuits, 1995
Traduction par Bernard Hoepffner

Chapitre 1	
Revue de la littérature	Le concept d'ambivalence La mesure de l'ambivalence
Etude préliminaire	La mesure des évaluations globales de consommateurs dans le e-commerce : pratiques et perception des internautes

↓

Chapitre 2	
Cadre conceptuel et hypothèses	Présentation de *l'Evaluative Space Grid* Résultats attendus

↓

Chapitre 3	
Etude 1	Validité convergente de *l'Evaluative Space Grid* avec les paires de continuums unipolaires
Etude 2	Validité convergente de *l'Evaluative Space Grid* avec les différentiels sémantiques, les discours des consommateurs et une mesure de l'ambivalence subjective

↓

Chapitre 4	
Etude 3	Test de *l'Evaluative Space Grid* comme outil de mesure de la satisfaction client

↓

Chapitre 5	
Etude 4	Etude qualitative après d'un échantillon d'experts : - qu'en pensent les chercheurs ? - qu'en pensent les professionnels ?

↓

Chapitre 6	
Etude 5	Place de l'incertitude dans les évaluations effectuées avec *l'Evaluative Space Grid*

– Chapitre 6 –

La place de l'incertitude dans les evaluations effectuées avec l'*Evaluative Space Grid*[12]

Nous avons jusqu'ici positionné la grille comme un outil permettant de séparer l'indifférence de l'ambivalence, deux réactions trop souvent agrégées au niveau du point central des différentiels sémantiques. Toutefois, les conclusions de l'étude qualitative incitent d'une part à établir clairement comment se répartissent ces deux types d'opinion le long de la diagonale de la grille de Larsen et d'autre part à intégrer la question des réponses incertaines à notre réflexion.

Notre objectif est de valider quantitativement les deux propositions principales mises en évidence par les chercheurs interrogés lors de l'étude qualitative (étude 4, Chapitre 5) concernant l'interprétation des différentes zones de la grille de Larsen. Premièrement, la diagonale de la grille, allant du point « pas du tout insatisfait ; pas du tout satisfait » au point « extrêmement insatisfait ; extrêmement satisfait » a été associée à la « charge émotionnelle » ou plus largement à l'intensité des réactions lors de l'évaluation. En d'autres termes, les participants ont considéré que plus on monte le long de la diagonale, plus

[12] *Cette étude a bénéficié du support de la Chaire AFNOR Performance des organisations de la Fondation Paris-Dauphine.*

l'évaluation est associée à des réactions fortes. Certains d'entre eux ont traduit cela spontanément comme une forme d'implication à l'égard de l'objet évalué. Le premier objectif de cette étude est de tester cette proposition.

Deuxièmement, le point central de la grille a été associé à la « volonté de ne pas se prononcer ». Plus précisément, celle-ci semble résulter de deux moteurs distincts dans le discours des experts. D'une part, certains participants ont expliqué avoir choisi ce point parce que leur opinion était réellement partagée entre des réactions positives et des réactions négatives. Ils employaient alors des expressions telles que « *verre à moitié vide, à moitié plein* », expérience « *pas wahou* » mais « *pas nulle non plus* ». On retrouve ici, l'idée selon laquelle le point central permet de saisir les évaluations ambivalentes, conformément aux travaux de Larsen et ses collègues (2009). D'autre part, le choix du point central pourrait traduire le fait que le répondant n'a « *pas un avis assez tranché* » pour réussir à « *se situer sur la question* ». Autrement dit, il « *n'arrive pas réellement à dégager de tendance claire* ». Cette idée laisse supposer que les réponses incertaines pourraient venir se « *réfugier* » au niveau du point central de la grille. Le deuxième objectif de cette étude est de tester cette proposition.

Section 1. Cadre théorique de l'étude

Dans la littérature, la distinction conceptuelle entre l'indifférence et l'ambivalence s'établit sur la base de certains critères. Ils peuvent être utilisés pour formuler des hypothèses sur la manière dont ces deux types d'évaluation devraient se répartir au sein de la grille.

1. Un premier critère de distinction entre indifférence et ambivalence : le degré d'intensité des réactions

Pour définir l'ambivalence, Conner et Sparks (2002) l'opposent à l'indifférence suivant un critère d'intensité. L'intensité permet de qualifier la puissance d'une réaction qui peut être, par exemple, très forte, modérée ou très faible (Zanna et Rempel, 1988). Suivant ce critère, l'ambivalence est composée d'un mélange de réactions positives et négatives d'intensité modérée à forte, tandis que l'indifférence se caractérise par une absence de réaction positive ou négative ou bien par des réactions de très faible intensité. Dans cette perspective, l'analyse effectuée par les participants de l'étude qualitative, consistant à associer la diagonale de la grille aux réactions émotionnelles prend tout son sens. Le

concept d'implication, c'est-à-dire le niveau d'intérêt d'un consommateur pour un produit ou un service, a également été mobilisé par les participants. Ce concept est intéressant parce qu'il permet d'inclure à la fois l'intensification des réactions affectives et cognitives. En effet, comme le soulignait Louisa, cette implication peut être « *émotionnelle ou ça peutêtre rationnel* ». Il s'agit par ailleurs d'un concept pertinent dans le cadre d'une utilisation de la grille dans un contexte de consommation.

2. Un critère à prendre en compte pour affiner cette distinction : le degré de certitude

2.1. Qu'est-ce qu'une évaluation incertaine ?

Le Grand Larousse 2014 définit l'incertitude comme « ce qui ne peut être établi avec exactitude ; doute ». Pour le *Petit Robert*[13], « être dans l'incertitude », c'est balancer, hésiter. L'incertain est quelque chose « dont la forme, la nature n'est pas nette, pas claire ». C'est « changeant, confus, indécis, indéfini, indéfinissable, imprécis, obscur, vague, flou ». Ces définitions généralistes peuvent être complétées par celle du dictionnaire de l'APA précisant que l'incertitude correspond à « un manque de confiance ou de clarté dans ses idées, ses décisions ou ses intentions ». Au contraire, la certitude d'une attitude correspond au « niveau de confiance subjective qu'une personne a concernant la validité de son attitude ». On retrouve ici la définition de Gross, Holtz et Miller (1995) pour lesquels il s'agit de la conviction avec laquelle on croit en son évaluation. De manière réciproque, Klopfer et Madden (1980) décrivent une attitude incertaine comme une réponse pour laquelle l'individu n'est pas capable d'identifier ses sentiments.

2.2. Quels liens peut-on établir entre incertitude, indifférence et ambivalence ?

Un premier élément de réponse se trouve peut-être dans les travaux de Suchman (1950). Il met en évidence une relation en « U » entre la certitude et la polarisation de l'attitude. Ainsi, les valeurs très positives et très négatives, sont associées à une certitude élevée tandis que les valeurs situées au niveau du centre du continuum allant du très négatif au très positif sont incertaines. Cette courbe en « U » a été confirmée dans les études d'Antil (1983), ainsi que dans les travaux de Sample et Warland (1973). Autrement dit, sur un

[13] *Ouvrage Collectif (2013), Le Petit Robert, Editions Le Robert, Paris.*

différentiel sémantique, les réponses situées au niveau du point central sont associées à un niveau d'incertitude élevé. Dans la mesure où les réactions indifférentes et ambivalentes sont habituellement agrégées à cet endroit du continuum, il convient de se demander si elles sont analysées comme des réponses incertaines dans la littérature.

Peut-on être certain que l'on est ambivalent ?

Après des décennies de recherches durant lesquelles l'ambivalence a été considérée comme une évaluation incertaine (Bassili, 1996 ; Gross, Holtz et Miller, 1995 ; Jonas, Diehl et Brömer, 1997 ; Petrocelli, Tormala et Rucker, 2007), cette idée tend à être remise en question. Le débat a été ouvert par Olsen (1999), soulignant que l'ambivalence pourrait être une réaction relativement certaine dans le cas de la cooccurrence de réactions positives et négatives de forte intensité. Autrement dit, il serait possible, dans certains cas, qu'un individu soit sûr de ressentir à la fois de fortes réactions positives et de fortes réactions négatives à l'égard d'un objet. Cette hypothèse, a été validée dans les travaux de Clarkson et ses collègues (2008) ou encore dans ceux de Petrocelli et ses collègues (2007). Ils ont démontré empiriquement que l'ambivalence et la certitude sont des concepts distincts. Si l'on pense par exemple à l'évaluation d'un service de commerce en ligne, un client peut à la fois être certain qu'il est satisfait de la livraison et certain qu'il est insatisfait du produit reçu. Dans cette perspective, il est donc possible d'être certain que l'on éprouve une attitude ambivalente à l'égard d'un objet.

Plus précisément, Heuvinck (2012) a établi une différence entre deux formes d'ambivalence : l'ambivalence manifeste et l'ambivalence anticipée. Les réactions d'ambivalence manifeste correspondent à des situations dans lesquelles les personnes sont capables d'identifier clairement les points positifs et négatifs qui composent leur évaluation ambivalente. En revanche, l'ambivalence anticipée fait référence au fait d'avoir à l'esprit de manière consciente un certain nombre d'informations positives (resp. négatives) à l'égard de l'objet évalué et d'anticiper d'éventuelles informations négatives (resp. positives).

Peut-on être certain que l'on est indifférent ?

A notre connaissance il n'existe pas de recherche établissant de relation entre l'indifférence et l'incertitude. Toutefois, il semble possible de transposer le raisonnement permettant d'expliquer le lien (in)certitude-ambivalence pour établir une relation entre l'(in)certitude et l'indifférence. S'il est possible d'être certain que l'on éprouve à la fois des réactions positives

et négatives à l'égard d'un objet, il semble également possible d'être sûr que l'on n'éprouve ni réaction positive, ni réaction négative à l'égard de celui-ci. La distinction effectuée par Nathalie dans l'étude qualitative va dans ce sens : « *"indifférent" veut dire deux choses […], tu peux avoir des croyances fortes par exemple à l'égard de Domenech, mais que ces croyances n'aient pas de charge émotionnelle. Par exemple, c'est un mauvais entraineur, pour moi être un mauvais entraineur, ça n'a pas de valeur émotionnelle, pour mon mari, être un mauvais entraîneur c'est hyper important […] en fait ça veut aussi bien dire pas d'émotion, cette personne est indifférente, et ça veut aussi dire je ne pense rien* ». Si l'on relit cet extrait en s'interrogeant sur le niveau de certitude des deux types d'indifférence que Nathalie a mis en évidence, on retrouve la même distinction que celle effectuée par Heuvinck (2012) concernant l'ambivalence. En effet, elle explique de manière très affirmative que pour elle, l'information « Domenech est un mauvais entraineur » n'est pas pertinente parce qu'elle ne s'intéresse pas au football. On peut donc penser qu'elle est certaine de son indifférence à l'égard des qualités de coach de Raymond Domenech. Il s'agirait dans ce cas d'une indifférence manifeste. Au contraire, l'indifférence associée au fait de ne pas avoir d'avis semble moins certaine. La rhétorique consistant à affirmer ne pas avoir d'avis sur une question dissimule sans doute un certain nombre de situations dans lesquelles les individus sont incertains de leur jugement. Il pourrait alors s'agir d'une indifférence anticipée, caractérisée par des opinions peu marquées et incertaines dans la mesure où l'individu anticipe qu'elles pourraient évoluer dans un sens ou dans l'autre.

3. Combinaison des critères pour expliquer la répartition des évaluations sur la diagonale de la grille

Nous avons distingué dans le premier point l'indifférence de l'ambivalence suivant un critère d'intensité des réactions. Le deuxième point a permis de mettre en évidence deux formes d'ambivalence et deux formes d'indifférence en fonction du degré de certitude qui leur est associé : une certitude forte pour les réactions manifestes d'indifférence ou d'ambivalence et une certitude faible pour les réactions d'indifférence et d'ambivalence anticipée. La combinaison de ces deux critères permet de mettre en évidence quatre types d'évaluations médianes dont l'articulation est décrite dans le tableau 32.

		Intensité	
		Faible	*Forte*
Certitude	*Faible*	Indifférence anticipée	Ambivalence anticipée
	Forte	Indifférence manifeste	Ambivalence manifeste

Tableau 33. Distinction entre indifférence et ambivalence suivant les critères d'intensité des réactions et de certitude dans la réponse donnée

Il s'agit maintenant de réfléchir à la localisation de ces différents types d'évaluation sur la diagonale de la grille. La phase qualitative de cette thèse a permis de faire émerger deux hypothèses. La première consiste à mettre à l'épreuve la distinction entre indifférence et ambivalence sous l'angle de l'implication, concept marketing pouvant traduire l'intensité des réactions. Ainsi le niveau d'implication devrait être croissant le long de la diagonale de la grille :

> *H1 : Il existe une relation linéaire positive entre les notes situées sur la diagonale de la matrice et le niveau d'implication à l'égard de l'objet évalué.*

La deuxième hypothèse vise à tester l'idée selon laquelle le point central de la grille correspond aux évaluations les moins certaines de la diagonale. Dans cette perspective, nous pensons que les évaluations anticipées, indifférentes ou ambivalentes, devraient s'agréger au niveau de ce point. Au contraire, les zones inférieure gauche et supérieure droite devraient être associées à un niveau certitude supérieur à celui du point médian.

> H2 : *Il existe une relation en « U » entre les notes situées sur la diagonale de la matrice et le niveau de certitude associé à celles-ci : les zones inférieure gauche et supérieure droite correspondent à des évaluations plus certaines que les réponses situées au niveau du point central.*

Pour tester ces hypothèses, nous réalisons une étude quantitative prenant la forme d'une enquête de satisfaction.

Section 2. Méthodologie

1. Choix de l'objet d'étude

Cette étude porte sur la satisfaction à l'égard du médecin généraliste. Comme nous l'avons déjà expliqué, Oliver (1997) souligne que les services présentent des propriétés propices à l'étude de l'ambivalence dans la mesure où il s'agit d'objets d'évaluations complexes, souvent composés d'une multitude d'attributs. C'est le cas de l'évaluation du médecin généraliste qui peut porter notamment, sur les qualités relationnelles du praticien, ses compétences médicales, sa propension à prescrire ce que le patient désire ou encore sur le confort du cabinet. La prise en compte de ces différents éléments peut se traduire par une évaluation composée à la fois de réactions positives et de réactions négatives, propice à l'utilisation de la grille.

2. Procédure et participants

Les données ont été collectées par un institut d'études, Easy Panel, dans le cadre d'une étude administrée en ligne. Ce recueil de données a été possible grâce au support financier de la Chaire AFNOR Performance des organisations de la Fondation Paris-Dauphine. Parmi les panélistes sollicités, 200 ont dûment rempli le questionnaire en contrepartie de points échangeables contre des cadeaux. Tous les répondants avaient entre 30 et 40 ans, tranche d'âge susceptible de garantir l'implication personnelle des participants dans une relation avec un médecin généraliste (M = 35,6 ans). L'échantillon était composé de 50% d'hommes et de femmes.

3. Format du questionnaire

L'étude se présentait comme une enquête de satisfaction classique. Elle commençait par quelques questions sur la durée de la relation avec le médecin généraliste et l'implication (Zaichkowsky, 1985) à l'égard du praticien. La mesure de la satisfaction globale à vis-à-vis du médecin généraliste arrivait en page suivante. Elle était réalisée à l'aide d'une échelle multi-items composée de 3 paires d'adjectifs antonymes fréquemment utilisés pour mesurer la satisfaction : « extrême insatisfaction ; extrême satisfaction » ; « opinions extrêmement négatives ; opinions extrêmement positives » et « opinions extrêmement défavorables ; opinions extrêmement favorables » (Anderson et Srinivasan, 2003 ; Lam &

al., 2004 ; Mittal, Ross et Baldasare, 1998). Ces trois items étaient mesurés avec des *Evaluative Space Grids*. Une brève explication d'utilisation de l'outil précédait la présentation de la première grille de Larsen. Chaque grille était accompagnée d'une mesure de la certitude dans la réponse donnée évaluée à l'aide de l'item suivant adapté de Petrocelli et collègues (2007) : « Pensez-vous que cette réponse reflète parfaitement votre opinion ? ». Une version adaptée du *Patient Satisfaction Questionnaire* (Grant et Hays, 1994) suivait ensuite dans le but de conforter les répondants dans l'idée qu'il s'agissait bien d'une étude de satisfaction. L'intégralité du questionnaire est présentée en annexe 6.

Section 3. Analyses et résultats

1. Codage préalable des données

Notre objectif était d'établir comment se répartissent les différents types d'évaluation sur la diagonale de la grille. Les analyses ont donc été focalisées sur les trois items de satisfaction globale mesurés avec des grilles de Larsen. Plus précisément, seules les réponses des participants ayant coché un point situé sur la diagonale de la grille pour chacun des trois items de mesure de la satisfaction globale ont été conservées. 50 observations sur 200 demeuraient à l'issue de cette procédure de sélection. Les points de la diagonale de la grille ont été codés de 1 (= « pas du tout satisfait » ; « pas du tout insatisfait ») à 5 (« extrêmement satisfait » ; « extrêmement insatisfait »). Suivant cette logique, le point central de la grille (« modérément insatisfait » ; « modérément satisfait ») a été codé 3.

2. Analyse des données

2.1. Dispersion des données

Pour commencer, nous avons étudié la dispersion des données sur la diagonale de la grille. Avec un alpha de Cronbach de 0,97, les trois items de mesure de la satisfaction globale offraient une bonne fiabilité. Parmi les 50 observations situées sur la diagonale, 23 ont été codées 3, ce qui signifie que le point central a capté près de la moitié des observations.

2.2. Test de l'hypothèse H1

Pour mettre à l'épreuve l'hypothèse H1, nous avons testé l'existence potentielle d'un lien entre les notes de la diagonale et le niveau d'implication à l'égard du médecin généraliste. Les résultats mettent en évidence une relation linéaire et positive ($p = 0{,}003$; $R^2 = 0{,}17$) entre les variables : plus la moyenne des notes situées sur la diagonale des grilles était élevée, plus l'implication moyenne à l'égard du médecin généraliste était forte. L'hypothèse H1 est donc validée

Toutefois, la faiblesse du R^2 laisse penser qu'il pouvait y avoir d'autres facteurs explicatifs de la relation entre les variables, ou que celle-ci n'est pas nécessairement linéaire. Nous avons exploré cette seconde piste à l'aide de l'approche proposée par Schoormans et Robben (1997). Il s'agit de comparer différents modèles de régression (linéaire, logarithmique, inverse, quadratique, cubique, exponentielle, logistique…) pour vérifier lequel ajuste le mieux le nuage de points formé par les valeurs des variables. La première étape consiste à observer le nuage de points (ou diagramme de dispersion), ainsi que les droites et courbes d'ajustements proposés par les différents modèles de régression. Une fonction du logiciel SPSS, intitulée « ajustement des fonctions », permet de comparer simultanément les ajustements proposés par les différents modèles.

Les résultats confirment nos intuitions. C'est le modèle cubique qui permet d'obtenir le R^2 le plus élevé ($p = 0{,}003$; $R^2 = 0{,}26$).

Dans la mesure où les fonctions cubiques sont toujours ascendantes, ce résultat va également dans le sens de l'hypothèse H1. En effet, cela signifie que dans tous les cas que les individus les moins impliqués ont choisi les notes basses de la diagonale et les individus les plus impliqués les notes hautes. Néanmoins, cette fonction nous indique qu'un creux se forme au centre gauche du nuage de points, soulignant un niveau d'implication plus faible au niveau du point central et du point qui le précède sur la diagonale de la grille. La figure 34 présentée en page suivante illustre cette analyse, révélant aussi que l'ajustement reste très imparfait, avec un R^2 faible même pour le modèle de régression cubique.

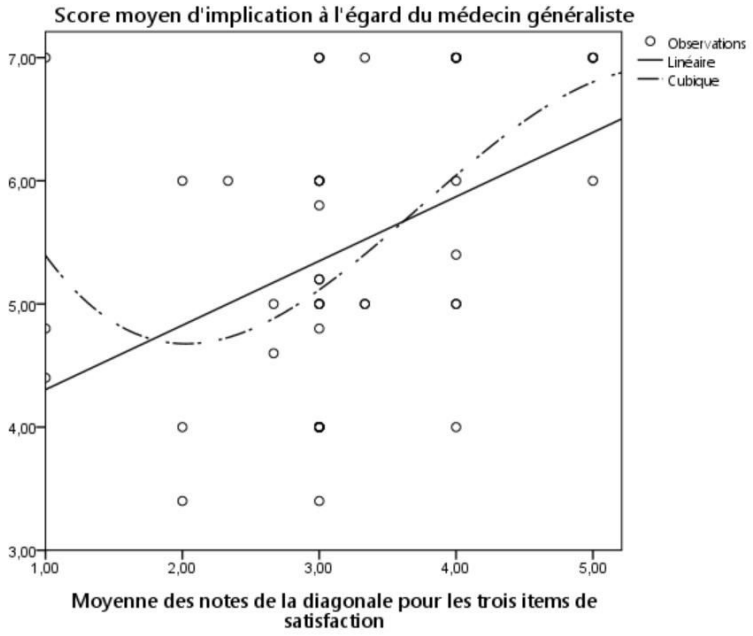


Figure 34. Ajustement de la relation entre les notes situées sur la diagonale de la grille et le niveau d'implication à l'égard du médecin suivant un modèle de régression linéaire et cubique

2.3. Test de l'hypothèse H2

Afin de montrer l'existence d'une relation curvilinéaire entre les notes situées sur la diagonale de la matrice et le degré de certitude qui leur est associé, nous avons à nouveau utilisé l'approche proposée par Schoormans et Robben (1997) et la commande « ajustement de fonctions » de SPSS. Nous cherchions cette fois-ci à comparer plus précisément les modèles de régression linéaire et quadratique pour vérifier lequel ajuste le mieux le nuage de points formé par les valeurs des variables.

Les résultats indiquent que c'est le modèle de régression quadratique qui décrit le mieux la relation. En effet, le modèle de régression linéaire est associé à un R^2 de 0,07 ($p = 0,06$), alors que le R^2 est de 0,40 ($p < 0,001$) pour le modèle quadratique. De plus, le coefficient de premier degré est négatif (-1,99) tandis que celui de second degré positif (0,35), signifiant que la relation suit une courbe en « U ». L'hypothèse H2 est validée (figure 35).

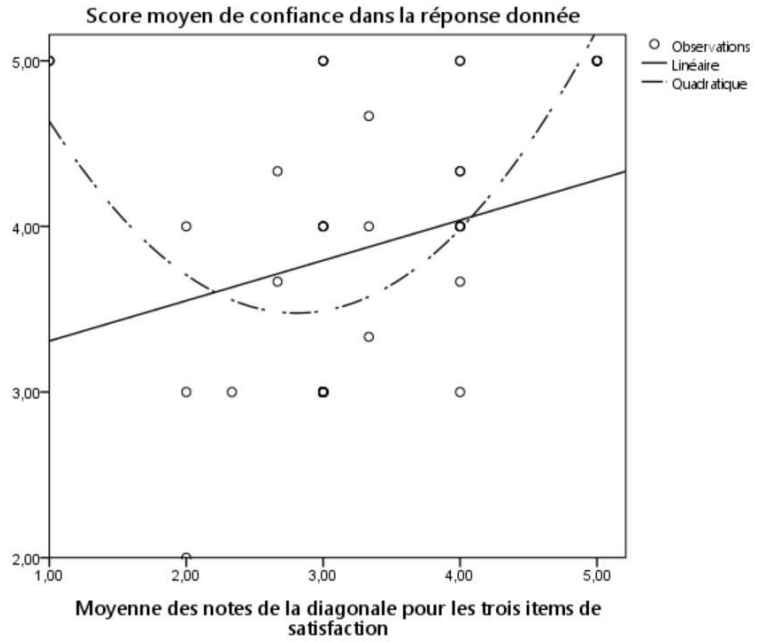


Figure 35. Ajustement de la relation entre les notes situées sur la diagonale de la grille et la confiance qui leur est associée suivant un modèle de régression linéaire et quadratique

Section 4. Conclusion

Cette étude permet de valider quantitativement les propositions issues de la phase qualitative de la thèse. Ainsi la diagonale de la matrice peut être mise en lien avec le niveau d'implication à l'égard de l'objet évalué et le degré de certitude dans l'évaluation donnée. Les variations du niveau d'implication permettent de distinguer deux types d'évaluations situées sur la diagonale de la grille : les valeurs basses de la diagonale (faible insatisfaction et faible satisfaction) sont caractérisées par un faible niveau d'implication tandis que les valeurs hautes de la diagonale (forte positivité et forte négativité) sont caractérisées par un niveau d'implication plus élevé. Toutefois les notes situées au niveau du point central et du point qui le précède sur la diagonale sont associées à un niveau d'implication plus faible. Ces résultats confirment la littérature indiquant que le degré d'intensité des réactions, ici mesuré par un score d'implication, permet de distinguer les évaluations indifférentes des évaluations ambivalentes. Autrement dit, sur la grille de Larsen les notes indifférentes sont plutôt situées en bas de la diagonale et les notes ambivalentes en haut de la diagonale. L'emplacement de la baisse d'implication le long de la diagonale est intéressant. A la lumière des résultats de l'étude préliminaire (cf. Chapitre 1, Section 3), on peut penser que ce relâchement correspond au flou sémantique identifié au niveau du centre gauche des

échelles composées de 5 étoiles. Dans cette perspective, la baisse d'implication se situerait au niveau du point de passage entre le négatif et le positif. D'autre part, en établissant une relation curvilinéaire entre les notes de la diagonale et le degré de certitude dans la note donnée, nous mettons clairement en évidence un point d'incertitude au centre de la grille.

Ces deux résultats permettent de valider le cadre d'analyse proposé suivant lequel les évaluations indifférentes et ambivalentes se répartissent le long de la diagonale de la grille en fonction des critères d'intensité des réactions et de certitude dans la note donnée. Plus précisément, dans la mesure où Heuvinck (2012) a mis en évidence le fait que l'ambivalence manifeste est associée à un niveau de certitude plus élevé que l'ambivalence anticipée, nous proposons la répartition suivante le long de la diagonale de la grille de Larsen :

- Les notes basses de la diagonale correspondent à de l'indifférence manifeste (faible intensité des réactions ; forte certitude dans la réponse donnée) ;
- Le point central de la diagonale permet de saisir les réactions anticipées d'indifférence et d'ambivalence (forte incertitude dans la réponse donnée),
- Les notes hautes de la diagonale correspondent à de l'ambivalence manifeste (forte intensité des réactions et forte certitude dans la réponse donnée).

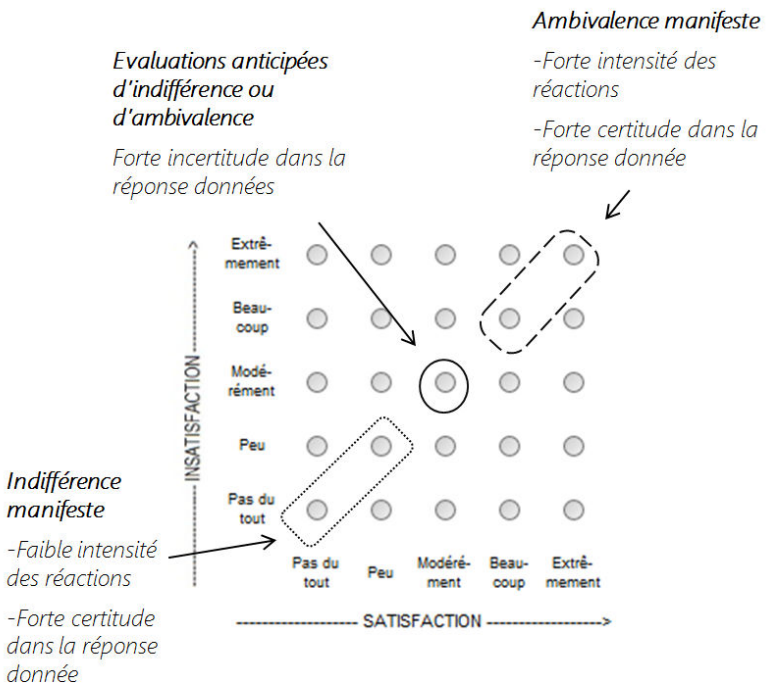


Figure 36. Répartition des évaluations médianes le long de la diagonale de l'*Evaluative Space Grid*

La figure 36 représente graphiquement la répartition des différents types d'évaluation suivant le cadre théorique proposé.

Section 5. Limites

Cette recherche soulève cependant certaines limites

Tout d'abord, malgré un échantillon initial de 200 personnes, la nécessité de sélectionner uniquement les observations situées sur la diagonale de la grille nous a conduits à réduire considérablement l'échantillon effectif. Parmi les observations restantes certains points de la diagonale n'ont été que rarement choisis, ce qui limite la portée des résultats.

Ensuite, le critère de l'implication ne semble pas être le seul critère explicatif de la répartition des répondants le long de la diagonale de la grille de Larsen. En effet, même avec le modèle de régression cubique, le R^2 restait faible : seule 26% de la variance des notes le long de la diagonale a pu être expliquée par le degré d'intérêt du répondant pour le service évalué. Cela soulève la nécessité de proposer des explications complémentaires ou alternatives.

D'autre part, la répartition des différents types d'évaluations médianes le long de la diagonale a simplement été inférée à *posteriori* de la grille de Larsen en cherchant à effectuer des liens avec d'autres variables. Par exemple, dans la mesure où elles étaient associées à un degré de certitude plus élevé et à une implication plus forte, nous avons supposé que les réponses situées en haut de la diagonale de la grille correspondaient à des réactions d'ambivalence manifeste. Nos conclusions pourraient être renforcées à l'aide d'un dispositif méthodologique expérimental manipulant *a priori* les différents types de réponses médianes dans le but d'observer directement où elles se répartissent sur la diagonale de l'*Evaluative Space Grid*.

***

– Conclusion générale –

La mesure des évaluations globales s'effectue classiquement sur un continuum linéaire, le différentiel sémantique. Cet outil est associé à un problème de mesure bien connu des chercheurs comme des professionnels des études : il agrège en son point central les évaluations ambivalentes, composées à la fois de réactions positives et négatives, et les évaluations indifférentes qui correspondent au contraire à l'absence ou à de très faibles de réactions positives ou négatives (Kaplan, 1972 ; Olsen, 1999 ; Thompson, Zanna et Griffin, 1995). La présente thèse montre la pertinence d'un nouvel outil de mesure, *l'Evaluative Space Grid* (Larsen & al., 2009), pour isoler la part d'ambivalence contenue dans les évaluations globales des consommateurs. Cette grille bidimensionnelle mesure séparément les réactions positives et les réactions négatives, et permet donc de distinguer quatre types d'évaluation de consommateurs :

- Les évaluations globalement positives, correspondant à des réactions positives fortes et des réactions négatives faibles ou inexistantes ;

- Les évaluations globalement négatives, correspondant à des réactions négatives fortes et des réactions positives faibles ou inexistantes ;

- Les évaluations globalement indifférentes, correspondant à des réactions positives faibles et des réactions négatives faibles ;

- Les évaluations globalement ambivalentes, correspondant à des réactions positives modérées à fortes et des réactions négatives modérées à fortes.

Cinq études empiriques ont été réalisées pour explorer comment cet outil de mesure pourrait trouver sa place dans la conduite d'études et de recherches en marketing. Les résultats de la recherche fournissent des contributions de nature méthodologique, théorique et managériale (Section 1), ainsi que des voies de recherche prometteuses (Section 2).

Section 1. Contributions

1. Contributions méthodologiques

La thèse apporte avant tout une contribution méthodologique à la littérature en marketing en y important un outil de mesure développé en psychologie qui permet de résoudre le problème du point central propre aux différentiels sémantiques. Plus qu'une simple transposition d'une discipline à une autre, notre travail explore comment cet outil de mesure pourrait trouver sa place dans la conduite d'études consommateurs.

Nous avons tout d'abord étendu les résultats de Larsen et ses collègues (2009) en montrant que l'*Evaluative Space Grid* permet effectivement, dans le cadre d'une application au champ du marketing, de distinguer les consommateurs indifférents des consommateurs ambivalents.

Les collectes de données réalisées à l'occasion de la thèse nous ont également donné l'occasion de tester différentes manières de présenter l'outil et son utilisation dans la cadre de questionnaires en ligne. Au fil des recueils de données, deux éléments sont apparus comme essentiels pour rendre l'outil compréhensible. D'une part, il est nécessaire de préciser qu'il s'agit de mesurer des évaluations globales composées d'une multitude d'éléments, certains pouvant être perçus positivement, d'autres négativement. D'autre part, l'ajout d'exemples d'évaluation illustrant notamment des situations d'ambivalence améliore grandement la prise en main de l'outil.

Notre travail apporte également une réponse à la crainte formulée par Larsen et ses collègues (2009) que seuls des étudiants présentant un niveau d'éducation élevé soient en mesure de s'approprier l'*Evaluative Space Grid*. Nous avons mis à l'épreuve cette hypothèse

en reprenant les données de l'étude 5 (Chapitre 4), collectées auprès de 200 panélistes de la société Easy Panel. En effet, dans cette base de données le niveau d'éducation des répondants variait du BEP-CAP au doctorat. Un test de comparaison de moyennes a montré qu'il n'y a pas de différence significative sur la variable « niveau de confiance dans la réponse donnée » en fonction du niveau d'éducation. Il semble donc que tous les profils de répondants soient en mesure de s'approprier l'*Evaluative Space Grid*.

Nous avons également montré que les réponses, pour lesquelles le répondant se déclare incertain se regroupent au centre de l'outil proposé, qu'il s'agisse de la grille de Larsen ou du différentiel sémantiques. Toutefois, à la différence du différentiel sémantique, l'*Evaluative Space Grid* permet de distinguer les réponses incertaines des réponses indifférentes et ambivalentes, et donc d'isoler les réponses à partir desquelles il est dangereux de tirer des conclusions.

2. Contributions théoriques

Au niveau théorique, ce travail contribue à l'approche bidimensionnelle des réactions évaluatives en intégrant le concept d'ambivalence au spectre des évaluations globales de consommateurs. En effet, si la littérature en psychologie comme en gestion reconnaît la nécessité d'envisager séparément les réactions positives et les réactions négatives, cela n'a jusqu'ici jamais été mis en lien avec le concept d'ambivalence des consommateurs. Nous proposons donc une nouvelle conception des évaluations globales de consommateurs impliquant quatre types de réactions : les réactions positives, négatives, indifférentes et ambivalentes.

Nous introduisons également l'incertitude à ce cadre d'analyse. Plus précisément, nous confirmons la distinction établie par Heuvinck (2012) entre ambivalence manifeste et ambivalence anticipée. Selon cet auteur, l'ambivalence manifeste correspond à des situations dans lesquelles l'individu peut identifier avec certitude les points positifs et les points négatifs qui caractérisent son ambivalence. En revanche, l'ambivalence anticipée décrit des situations dans lesquelles l'individu identifie clairement un certain nombre d'informations positives (resp. négatives) à l'égard de l'objet évalué, mais anticipe en même temps d'éventuelles informations négatives (resp. positives). En conséquence, l'ambivalence anticipée intègre une part d'incertitude et c'est précisément ce qui la distingue de l'ambivalence manifeste.

Nous allons plus loin en proposant une distinction entre indifférence manifeste et indifférence anticipée. L'indifférence manifeste correspond à une absence avérée de réaction à l'égard de l'objet évalué. Elle est assumée voire revendiquée pour souligner un manque d'intérêt ressenti avec certitude. L'indifférence anticipée décrit, quant à elle, une absence d'avis beaucoup moins catégorique : la personne ne se prononce pas parce qu'elle n'est pas sûre de son jugement. Ainsi, nous proposons d'ajouter cette distinction entre indifférence manifeste et indifférence anticipée à celle proposée par Heuvinck (2012) entre ambivalence manifeste et ambivalence anticipée. En effet, les réactions anticipées, qu'elles soient d'ambivalence ou d'indifférence, sont avant tout des réactions incertaines. En cohérence avec notre approche méthodologique, nous proposons donc de distinguer au final trois types de réactions « médianes » : les réactions manifestement ambivalentes, les réactions manifestement indifférentes et les réactions incertaines.

3. Contributions managériales

D'un point de vue managérial, ce travail apporte une réponse aux praticiens désireux de comprendre ce qui se cache derrière des performances modérées. Contrairement à ce que préconise l'approche du *Net Promoter Score*, nous montrons qu'il ne faut pas éliminer sans distinction les répondants qualifiés de « passifs » par Reichheld (2003). Plus précisément, l'*Evaluative Space Grid* permet de faire le tri au sein de ces évaluations, entre celles qui sont informatives et celles qui ne le sont pas.

La première recommandation que nous adressons aux praticiens est d'exclure de leurs analyses les évaluations incertaines. En effet, même si elles correspondent à une part non négligeable des réponses, il semble risqué de tirer des conclusions à partir de ces réactions dans la mesure où elles sont moins prédictives des comportements (Bizer & al., 2006 ; Fazio et Zanna, 1978 ; Petrocelli, Tormala et Rucker, 2007 ; Rucker et Petty, 2004) et sont moins durables (Bassili, 1996).

Les praticiens doivent se concentrer sur les évaluations indifférentes et les évaluations ambivalentes, qui fournissent pour leur part une réelle information. Leur distinction peut, par exemple, être intéressante dans le cadre du pilotage d'une marque. En effet, dans ce contexte, les recommandations ne sont pas du tout les mêmes face à ces deux types d'évaluations : des nouveautés pourraient susciter l'intérêt des consommateurs indifférents,

alors que pour séduire les consommateurs ambivalents, il vaudrait mieux focaliser les efforts sur la levée des points d'insatisfaction qui composent leur évaluation.

En ce qui concerne la prédiction des comportements, les conclusions sont très claires. Reconsidérons par exemple les données collectées dans le cadre de l'étude 3 (Chapitre 4) sur la satisfaction à l'égard d'une banque, et notamment la moyenne d'intention de bouche-à-oreille positif. Cette moyenne s'élève à 4,5/7 chez les répondants ambivalents, à 3,8/7 chez les répondants indifférents (soit en dessous de la moyenne générale de 4,14/7) et à 4/7 chez les répondants incertains, indiquant que ces derniers ont eu tendance à choisir le point central sur le continuum. Plus précisément, la différence est significative entre les intentions de bouche-à-oreille des ambivalents par rapport aux indifférents et aux incertains, comme l'indique le tableau 33[14].

	Effectif	Moyenne
Insatisfaits	12	1,81
Indifférents	37	3,81[a]
Incertains	40	4[b]
Ambivalents	24	4,51[ab]
Satisfaits	39	5,09[ab]
Total / Moyenne	152	4,14

Les lettres a et b indiquent les groupes entre lesquels les moyennes diffèrent significativement (p<0,05)
Les résultats du test ne sont pas interprétés pour les insatisfaits compte tenu de la faiblesse de l'effectif

Tableau 34. Test de différence de moyennes d'intention de bouche-à-oreille positif en fonction du type d'évaluation globale

[14] *Les résultats sont similaires pour les intentions de fidélité, mais les écarts entre les groupes sont moins marqués.*

Section 2. Limites et voies de recherche

Les limites de chacune des 5 études que nous avons réalisées ont présentées spécifiquement à la fin de chaque chapitre. Cette section est consacrée à la discussion de limites transverses qui tiennent au dispositif de recherche général.

1. Vers d'autres opérationnalisations

Cette thèse s'est limitée au test d'une seule « opérationnalisation » de l'*Evaluative Space Grid*. Il s'agissait d'un parti pris assumé visant à rendre possible la comparaison des résultats des études entre eux. Ainsi, les réactions négatives ont systématiquement été mesurées sur l'axe des ordonnées, tandis que les réactions positives ont systématiquement été mesurées sur l'axe des abscisses. Chaque axe affichait 5 points, générant par conséquent systématiquement 25 modalités de réponses différentes. Enfin, l'origine de la grille (c'est-à-dire le point de coordonnées 0 ; 0) a systématiquement été placée en bas à gauche de l'outil. Ce premier travail d'exploration de l'intérêt de l'*Evaluative Space Grid* dans la conduite d'études consommateurs étant terminé, nous appelons à tester d'autres opérationnalisations de l'*Evaluative Space Grid*. En particulier, nous avons réfléchi à un certain nombre de modifications susceptibles de prolonger les recherches sur cet outil.

Tout d'abord, le nom des étiquettes sémantiques qualifiant les points des axes pourrait faire l'objet de recherches futures. Selon la variable mesurée, nous avons opté pour les termes « pas du tout » ou « aucune » pour qualifier le premier point des axes. Or, pour certains des experts interrogés dans le cadre de l'étude qualitative, ces étiquettes, lorsqu'elles sont placées sur l'axe des abscisses mesurant les réactions positives, ont toutefois pu être interprétées comme un début d'insatisfaction. Cette interprétation est gênante dans la mesure où les réactions négatives ou d'insatisfaction sont censées être mesurées sur l'axe des ordonnées. Il est donc possible que ce choix ait influencé les résultats. Pour le vérifier, il faudrait comparer l'opérationnalisation retenue dans la thèse à des opérationnalisations alternatives (par exemple graduer les axes de « très faible » à « très fort ») dans le cadre d'un design inter-sujets.

Se pose également la question du nombre total de points de l'outil. Lors de l'étude qualitative, certains participants ont semblé impressionnés par les 25 points de l'*Evaluative Space Grid*. Réduire chaque axe à 3 points permettrait de limiter la grille à 9 réponses possibles. Comme l'illustre la figure 37, les différentes modalités de réponses deviennent

alors très faciles à interpréter. Deux zones d'évaluations polarisées se dégagent en haut à gauche et en bas à droite. Chacune est composée de 3 points, ce qui permet d'exprimer des nuances. Par ailleurs, les 3 points de la diagonale sont suffisants pour isoler les réponses indifférentes, des réponses incertaines et des réponses ambivalentes. Cette piste de recherche parait donc particulièrement intéressante à poursuivre pour limiter le flou interprétatif associé à certaines modalités de réponse d'une grille comportant 25 points.

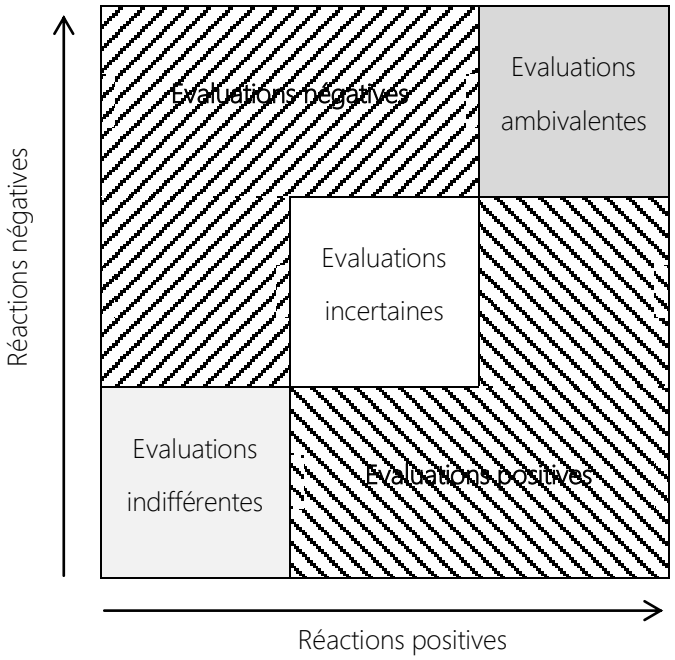


Figure 37. Répartition des zones évaluatives sur une *Evaluative Space Grid* en 9 points

Au-delà du nombre de points de l'outil, la présence même de points peut être discutée. La programmation de l'*Evaluative Space Grid* sous forme de « carte thermique » (possible par exemple avec le logiciel Qualtrics) permettrait de proposer aux répondants un carré vierge et la possibilité de cliquer à n'importe quel endroit de ce carré pour indiquer le niveau de leurs réactions positives et négatives. Cela offrirait la possibilité d'associer à chaque clic un couple de coordonnées sur la carte thermique et d'identifier de manière précise une infinité de combinaisons possibles entre les réactions positives et négatives exprimées par les individus. De telles données offriraient des perspectives nouvelles en termes d'analyse en passant d'une variable discrète à une variable continue. Sans aller jusqu'à la suppression totale des points, une variante consisterait à matérialiser physiquement les différentes zones évaluatives sur la grille avec des cadrans de couleur pour guider davantage les répondants comme cela est fait dans la figure 37.

Enfin, la question de l'emplacement de l'origine de l'*Evaluative Space Grid* mérite que l'on s'y attarde un peu. En effet, comme nous l'avons expliqué dans l'introduction de la partie empirique, nous avons fait le choix de situer le point de coordonnées [0 ; 0] en bas à gauche de la grille, alors que Larsen et ses collègues (2009) ont démontré que cette opérationnalisation peut augmenter le niveau des évaluations négatives lorsqu'il n'est pas possible de contraindre le curseur au niveau de l'origine de l'outil (situation dans laquelle nos collectes ont été effectuées). Nous avons fait ce choix dans le but d'inciter les répondants à prendre en compte la part de négativité contenue dans leur évaluation. L'objectif de cette démarche était d'apporter une solution au biais de positivité très souvent constaté dans la mesure de la satisfaction (Hurley et Estelami, 1998 ; Peterson et Wilson, 1992).

Pour estimer l'intérêt potentiel de cette piste de recherche, nous avons repris les données de l'étude 3 (Chapitre 4). Pour rappel, dans cette étude nous avions, mesuré la satisfaction globale à l'aide de trois items dans la cadre d'un design inter-sujets : la moitié des participants avaient effectué leur évaluation sur des grilles de Larsen et l'autre moitié sur des différentiels sémantiques. Dans le but d'obtenir des données comparables dans les deux conditions, pour l'ensemble de données collectées avec la grille de Larsen nous avons extrait les coordonnées des notes et retranché la note de négativité à la note de positivité. Cette opération permet de ramener les données de l'*Evaluative Space Grid* sur un plan univarié. Nous avons ensuite calculé des coefficients de dissymétrie pour nos deux ensembles de données. Ce coefficient varie de -1 à 1 et indique dans quelle mesure la distribution des données s'écarte d'une loi normale. Un coefficient positif décrit une distribution décalée à gauche de la médiane et une queue de distribution étalée vers la droite, tandis qu'un coefficient négatif révèle une distribution décalée à droite de la médiane et une queue de distribution étalée vers la gauche.

Le tableau 34 indique que dans les deux conditions expérimentales, presque tous les coefficients d'asymétrie calculés sont négatifs, comme s'est habituellement le cas pour la satisfaction. Toutefois, dans la condition *Evaluative Space Grid*, les coefficients sont beaucoup plus proches de zéro, soulignant que la distribution des données se rapproche davantage d'une distribution symétrique. Cette analyse confirme le fait que la grille de Larsen incite les répondants à prendre conscience des éléments négatifs latents qui composent leur évaluation. Face à ces résultats encourageant, nous invitons les futures recherches à se pencher davantage sur les perspectives qu'offre la grille de Larsen en

termes de réduction de la distribution asymétrique négative de la satisfaction. En effet, cela peut constituer un enjeu par exemple dans le cadre de tests de nouveaux concepts.

	Moyenne (écart type)		Asymétrie [erreur standard]	
	Condition DS	*Condition ESG*	*Condition DS*	*Condition ESG*
Opinions extrêmement négatives - Opinions extrêmement positives	6,10 (1,85)	5,41 (1,53)	-0,718 [0,19]	-,093 [0,20]
Extrême insatisfaction - Extrême satisfaction	6,08 (1,74)	5,50 (1,57)	-0,648 [0,19]	0,114 [0,20]
Opinions extrêmement défavorables - Opinions extrêmement favorables	6,06 (1,90)	5,59 (1,48)	-0,593 [0,19]	-0,015 [0,20]
Moyenne des trois items	6,08 (1,64)	5,50 (1,44)	-0,487 [0,19]	0,040 [0,20]

Tableau 35. Comparaison des coefficients d'asymétrie pour les items de mesure de la satisfaction globale en fonction de l'outil de mesure utilisé (différentiel sémantique ou *Evaluative Space Grid*)

2. Vers l'étude de variables individuelles

L'*Evaluative Space Grid* pourrait également servir de support à l'étude de diverses variables individuelles.

Tout d'abord, elle pourrait être utilisée pour identifier des consommateurs présentant une tendance chronique à l'ambivalence. En effet, Ito et Cacioppo (2005) ont démontré que les individus diffèrent dans la manière dont ils évaluent la positivité et la négativité d'un stimulus et que ces différences individuelles sont stables dans le temps. En d'autres termes, certaines personnes font l'objet d'une positivité ou d'une négativité dispositionnelle liée à un biais individuel du processus évaluatif : en moyenne elles évaluent plus positivement (resp. négativement) que les autres les stimuli positifs (resp. négatifs). Norris et ses collègues (2011) ont confirmé l'existence de ces différences individuelles dans le cadre de diverses tâches de jugements rapportées sur des *Evaluative Space Grids*. De la même manière, nous pensons que certaines personnes ont des difficultés récurrentes à intégrer les éléments positifs et négatifs qui composent leurs évaluations et sont confrontées plus souvent que les autres à des réactions ambivalentes. L'*Evaluative Space Grid* pourrait être utilisée pour identifier ces personnes en effectuant des mesures répétées à partir de stimuli normés. L'objectif serait alors de repérer les individus qui ont tendance à situer leurs évaluations dans le coin supérieur droit de la grille de Larsen. Des actions de

communication spécifiques pourraient être envisagées pour aider ces personnes à résoudre leur ambivalence en ciblant les freins qui entravent leur persuasion.

Enfin, nous avons postulé dans le paragraphe précédant que les individus éprouvaient le besoin de résoudre leurs contradictions, mais ce besoin apparaît très marqué culturellement. Pour illustration, l'héritage des penseurs grecs incite les Occidentaux à résoudre leurs contradictions : la loi aristotélicienne de la « liberté de contradiction » stipule qu'une chose n'est pas identique à son contraire et celle du « tiers exclu » précise que toute proposition est soit vraie soit fausse. Au contraire, la pensée dialectique orientale accepte les contradictions et va jusqu'à considérer qu'elles sont nécessaires à l'émergence de la vérité (Nisbett & al., 2001). Ces différences culturelles offrent des pistes de recherche intéressantes pour l'étude de l'ambivalence en général, mais également en ce qui concerne l'utilisation d'un outil tel que l'*Evaluative Space Grid* qui suppose d'accepter de rapporter des réactions contradictoires. On peut ainsi imaginer que les Orientaux s'approprieraient plus facilement ce nouvel outil que les Occidentaux, et qu'il permettrait de prédire plus efficacement leurs comportements de consommation.

– Références –

Anderson, E. W., Fornell, C. et Lehmann, D. R. (1994). Customer Satisfaction, Market Share, and Profitability: Findings from Sweden, *Journal of Marketing*, *58* (July), 53-66.

Anderson, R. E. et Srinivasan, S. C. (2003). E-Satisfaction and E-Loyalty: A Contingency Framework, *Psychology and Marketing*, *20* (2), 123-138.

Andrade, E. B. et Cohen, J. B. (2007). On the Consumption of Negative Feelings, *Journal of Consumer Research*, *24*, 283-300.

Antil, J. H. (1983). Uses of Response Certainty in Attitude Measurement, *Advances in Consumer Research*, *10*, 409-415.

Armitage, C. J. (2003). Beyond Attitudinal Ambivalence: Effects of Belief Homogeneity on Attitude-Intention-Behaviour Relations, *European Journal of Social Psychology*, *33*, 551-563.

Armitage, C. J. et Conner, M. (2000). Attitudinal Ambivalence: a Test of Three Key Hypotheses, *Personality and Social Psychology Bulletin*, *26*, 1421-1432.

Bargh, J. A., Chaiken, S., Govender, R. et Pratto, F. (1992). The Generality of the Automatic Attitude Activation Effect, *Journal of Personality and Social Psychology, 62*, 893–912.

Baron, R. M. et 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.

Bassili, J. N. (1996). Meta-Judgmental versus Operative Indexes of Psychological Attributes: The Case of Measures of Attitude Strength, *Journal of Personality and Social Psychology, 71*, 637–653.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C. et Vohs, K. D. (2001). Bad is Stronger than Good, *Review of General Psychology, 5*, 323-370.

Bell, D. W. et Esses, V. M. (2002). Ambivalence and Response Amplification: a Motivational Perspective, *Personality and Social Psychology Bulletin, 28*, 1143-1152.

Bickart, B. A. (1993). Carryover and Backfire Effects in Marketing Research, *Journal of Marketing, 39* (February), 53-66.

Bizer, G. Y., Tormala, Z. L., Rucker, D. D. et Petty, R. E. (2006). Memory-Based versus On-Line Processing: Implications for Attitude Strength, *Journal of Experimental Social Psychology, 42*, 646–653.

Bleuler, E. (1950). *Dementia Praecox or the Group of Schizophrenias*, New York, International University Press.

Breckler, S. J. (1994). A Comparison of Numerical Indices for Measuring Attitude Ambivalence, *Educational and Psychological Measurement, 54*, 350-365.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1-36), Hillsdale, NJ, Erlbaum.

Cacioppo, J. T. et Berntson, G. G. (1994). Relationship Between Attitudes and Evaluative Space: a Critical Review, with Emphasis on the Separability of Positive and Negative Substrates, *Psychological Bulletin, 115*, 401-423.

Cacioppo, J. T., Berntson, G. G., Norris, C. J. et Gollan, J. K. (2011). *The Evaluative Space Model*, Thousand Oaks, CA, Sage Press.

Cacioppo, J. T., Gardner, W. L. et Berntson, G. G. (1997). Beyong Bipolar Conceptualizations and Measures: the Case of Attitudes and Evaluative Space, *Personality and Social Psychology Review, 1*, 3-25.

Chandon, J.-L. et Bartikowski, B. (2004). Une échelle ordinale permettant de classer les répondants en « satisfait », « indifférent » et « insatisfait », *Recherche et Applications en Marketing, 9* (1), 39-53.

Clark, J. K., Wegener, D. T. et Fabrigar, L. R. (2008). Attitudinal Ambivalence and Message-Based Persuasion: Motivated Processing of Proattitudinal Information and Avoidance of Counterattitudinal Information *Personality and Social Psychology Bulletin, 34*, 565-577.

Clarkson, J. J., Tormala, Z. L. et Rucker, D. D. (2008). A New Look at the Consequences of Attitude Certainty: The Amplification Hypothesis, *Journal of Personality and Social Psychology, 95* (4), 810-825.

Conner, M. et Sparks, P. (2002). Ambivalence and Attitudes, *European Review of Social Psychology, 12*, 37–70.

Costarelli, S. et Colloca, P. (2004). The Effects of Attitudinal Ambivalence on Pro-environmental Behavioural Intentions, *Journal of Environmental Psychology, 24*, 279-288.

Creswell, J. W. (1995). *Research Design: Qualitative and quantitative Approaches*, CA: Thousand Oaks, Sage Press.

Cronin, J. J., Brady, M. K. et Hult, T. M. (2000). Assessing the Effects of Quality, Value, and Customer Satisfaction on Behavioral Intention in Service Environments?, *Journal of Retailing, 76* (2), 193-218.

Cronin Jr, J. J. et Taylor, S. A. (1992). Measuring service quality: a reexamination and extension, *The Journal of Marketing*, 55-68.

Darpy, D. (2012). *Comportements du consommateur : Concepts et outils*, Dunod.

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*, New York, McGraw-Hill.

Dubois, B. et Burns, J. A. (1975). An analysis of the Meaning of the Question Mark Response Category in Attitude Scales, *Journal of Educational and Psychological Measurement, 35*, 869-884.

Edwards, J. R. et Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis, *Psychological Methods*, *12*, 1-22.

Einhorn, H. (1970). The Use of Nonlinear, Noncompensatory Models in Decision Making, *Psychological Bulletin*, *73*, 221-230.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I. et Heider, K. (1987). Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion, *Journal of Personality and Social Psychology*, *53*, 712-717.

Estelami, H. et De Maeyer, P. (1997). A Visual Approach for Identifying Consumer Satisfaction Segments, *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, *10*, 104-115.

Evanschitzky, H. et Wunderlich, M. (2006). An Examination of Moderator Effects in the Four-Stage Loyalty Model, *Journal of Service Research*, *8* (4), 330-345.

Evrard, Y., Pras, B. et Roux, E. (2003). *Market : Etudes et recherche en marketing* (3ème ed.), Paris, Dunod.

Fazio, R. H. et Zanna, M. P. (1978). Attitudinal Qualities Relating to the Strength of the Attitude-Behavior Relationship *Journal of Experimental Social Psychology*, *14*, 398–408.

Fishbein, M. (1963). An Investigation of the Relationship Between Beliefs About an Object and the Attitude Toward that Object, *Human Relations*, *16*, 233-240.

Fishbein, M. (1967). Attitude and the Prediction of Behavior. In M. Fishbein (Ed.), *Readings in Attitude Theory and Measurement*, New York:, John Wiley & Sons, Inc.

Forget, M.-H. (2013). Le développement des méthodes de verbalisation de l'action : un apport certain à la recherche qualitative, *Recherches qualitatives*, *32* (1), 57-80.

Fornell, C. et Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error, *Journal of Marketing Research*, *18* (1), 39-50.

Ghiglione, R. et Matalon, B. (1978). *Les enquêtes sociologiques. Théories et pratique*, Paris, Armand Colin.

Graham, J. R. (1993). *MMPI-2: Assessing Personality and Psychopathology*, New York, Oxford University Press.

Grant, N. M. et Hays, R. D. (1994). *The Patient Satisfaction Questionnaire Short-Form (PQS-18)*, CA: Santa Monica, RAND.

Green, P. et Wind, Y. (1973). *Muttiattribute Decisions in Marketing*, Hinsdale, Illinois: , Dryden Press.

Greimas, A. J. et Courtès, J. (1993). *Sémiotique: Dictionnaire raisonné de la théorie du langage*, Paris, Hachette Supérieur.

Gross, S. R., Holtz, R. et Miller, N. (1995). Attitude certainty. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 215-246), Mahwah, NJ, Lawerence Erlbaum Association.

Hair, J., Black, W., Babin, B., Anderson, R. et Tatham, R. (2006). *Multivariate Data Analysis* (6 ed.), NJ: Pearson Prentice Hall, Upper Saddle River.

Herzberg, F., Mausner, B. et Snyderman, B. (1959). *The Motivation to Work*, New York, John Wiley and sons, Inc.

Heuvinck, N. (2012). *Underneath the Skin of Attitude Ambivalence: Different Types of Ambivalence*. Doctoral dissertation, University of Gent.

Hu, L. et Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification, *Psychological Methods*, *3* (4), 424-456.

Hurley, R. F. et Estelami, H. (1998). Alternative Indexes for Monitoring Customer Perceptions of Service Quality: a Comparative Evaluation in Retail Context, *Journal of the Academy of Marketing Science*, *26* (3), 209-221.

Ito, T. A. et Cacioppo, J. T. (2005). Variations on a Human Universal : Individual Differences in Positivity Offset and Negativity Bias, *Cognition and Emotion*, *19* (1), 1-26.

Jamieson, D. W. (1988). *The Influence of Value Conflicts on Attitudinal Ambivalence*. Paper presented at the Annual meeting of the Canadian Psychological Association, Montréal.

Jolibert, A. et Jourdan, P. (2005). *Marketing research*, Paris, Dunod.

Jonas, K., Brömer, P. et Diehl, M. (2000). Attitudinal Ambivalence, *European Review of Social Psychology*, *11*, 35-74.

Jonas, K., Diehl, M. et Brömer, P. (1997). Effects of Attitudinal Ambivalence on Information Processing and Attitude-Intention Consistency, *Journal of Experimental Social Psychology*, *33*, 190-210.

Judd, C. M. et Brauer, M. (1995). Repetition and Evaluative Extremity. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 43-72), Mahwah, NJ, Lawrence Erlbaum Association.

Kachadourian, L. K., Fincham, F. et Davila, J. (2005). Attitudinal Ambivalence, Rumination, and Forgiveness of Partner Transgressions in Marriage, *Personality and Social Psychology Bulletin* (31), 334-342.

Kaplan, K. J. (1972). On the Ambivalence-Indifference Problem in Attitude Theory and Measurement: a Suggested Modification of the Semantic Differential Technique, *Psychological Bulletin*, *77* (5), 361-372.

Karatepe, O. M., Yavas, U. et Babakus, E. (2005). Measuring Service Quality of Banks: Scale Development and Validation, *Journal of Retailing and Consumer Services*, *12* (5), 373-383.

Katz, D. (1944). The Measurement of Intensity. In H. Cantril (Ed.), *Gauging Public Opinion* (pp. 51-65), Princeton, NJ, Princeton University Press.

Katz, I. et Hass, R. G. (1988). Racial Ambivalence and American Value Conflict: Correlational and Priming Studies of Dual Cognitive Structures, *Journal of Personality and Social Psychology*, *55*, 893–905.

Kernberg, O. F. (1976). *Object-Relations Theory and Clinical Psychoanalysis*, New York, Simon & Schuster.

Klopfer, F. J. et Madden, T. M. (1980). The Middlemost Choice on Attitude Items: Ambivalence, Neutrality, or Uncertainty?, *Personality and Social Psychology Bulletin*, *6* (97-101).

Kopfler, F. J. et Madden, T. M. (1980). The Middlemost Choice on Attitude Items: Ambivalence, Neutrality, or Uncertainty?, *Personality and Social Psychology Bulletin*, *6* (97-101).

Krosnick, J. A. et Petty, R. E. (1995a). Attitude Strenght: an Overview. In J. A. Krosnick & R. E. Petty (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 1-124), Mahwah, N.J., Lawrence Erlbaum Associates.

Krosnick, J. A. et Petty, R. E. (1995b). *Attitude Strength: Antecedents and Consequences*, Mahwah, N.J., Lawrence Erlbaum Associates.

Kuusela, H. et Paul, P. (2000). A Comparison of Concurrent and Retrospective Verbal Protocol Analysis, *The American Journal of Psychology*, *113* (3), 387-404.

Lam, S. Y., Shankar, V., Erramilli, M. K. et Murthy, B. (2004). Customer Value, Satisfaction, Loyalty, and Switching Costs: An Illustration From a Business-to-Business Service Context, *Journal of the Academy of Marketing Science*, *32* (4), 293-311.

Larsen, J. T. et McGraw, A. P. (2011). Further Evidence for Mixed Emotions, *Journal of Personality and Social Psychology*, *100* (6), 1095–1110.

Larsen, J. T., McGraw, A. P. et Cacioppo, J. T. (2001). Can People Feel Happy and Sad at the same Time? , *Journal of Personality and Social Psychology*, *81* (4), 684-696.

Larsen, J. T., Norris, C. J., McGraw, A. P., Hawkley, L. C. et Cacioppo, J. T. (2009). The Evaluative Space Grid: A Single-Item Measure of Positivity and Negativity, *Cognition and Emotion*, *23* (3), 453-480.

Lavine, H., Borgida, H. et Sullivan, J. L. (2000). On the Relationship Between Attitude Involvement and Attitude Accessibility: Toward a Cognitive-Motivational Model of Political Information Processing, *Political Psychology*, *21* (1), 81-106.

Lehmann, D. (1971). Television Show Preferences: Application of a Choice Model, *Journal of Marketing Research*, *8* (February), 47-55.

Leygue, C. (2009). *Aspects structuraux des attitudes : Antécédents et conséquences de l'ambivalence ressentie.* Thèse de Doctorat, Université Paris-Ouest Nanterre La Défense.

Lichtlé, M. C., Llosa, S. et Plichon, V. (2002). La contribution des éléments d'une grande surface alimentaire à la satisfaction du client, *Recherche et Applications en Marketing*, *17* (4), 23-34.

Little, T. D., Lindenberger, U. et Nesselroade, J. R. (1999). On Selecting Indicators for Multivariate Measurement and Modeling with Latent Variables: When "Good"

Indicators are Bad and "Bad" Indicators are Good, *Psychological Methods*, *4* (2), 192-211.

Llosa, S. (1996). *Contributions à l'étude de la satisfaction dans les services.* Thèse de Doctorat, IAE d'Aix en Provence, Université Aix Marseille III.

MacKinnon, D. P., Lockwood, C. M. et Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods, *Multivariate Behavioral Research*, *39* (1), 99-128.

Maheswaran, D. et Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment, *Journal of Personality and Social Psychology*, *61* (1), 13.

Maio, G. R., Bell, D. W. et Esses, V. M. (1996). Ambivalence and Persuasion: the Processing of Messages about Immigrant Groups, *Journal of Experimental Social Psychology*, *32*, 513-536.

Matlin, L. et Stang, D. (1978). *The Pollyanna Principle Selectivity, in Language, Memory and Thought*, MA: Cambridge, Senkman Publishing Company.

McGregor, I., Newby-Clark, I. R. et Zanna, M. P. (1999). Remembering Dissonance: Simultaneous Accessibility of Inconsistent Cognitive Elements Moderates Epistemic Discomfort. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology* (pp. 325-353), Washington, DC, American Psychology Association.

Meehl, P. E. (1964). *Manual for Use with Checklist of Schizotypic Signs*. Unpublished Work. University of Minnesota Medical School, Minneapolis.

Miles, M. B. et Huberman, A. M. (1991). *Analyses de données qualitatives*, Bruxelles, De Boeck-Wesmael.

Mittal, V., Ross, W. T. J. et Baldasare, P. M. (1998). The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions, *Journal of Marketing*, *62* (1), 33-47.

Mucchielli, R. (1999). *Le questionnaire dans l'enquête psycho-sociale*, ESF éditeur.

Neapolitan, R. E. (2004). *Learning Bayesian Networks*, Upper Saddle River, NJ, Prentice Hall.

Newby-Clark, I. R., McGregor, I. et Zanna, M. P. (2002). Thinking and Caring about Cognitive Inconsistency: When and for Whom does Attitudinal Ambivalence Feel Uncomfortable?, *Journal of Personality and Social Psychology*, *82*, 157–166.

Nisbett, R. E., Peng, K., Choi, I. et Norenzayan, A. (2001). Culture and Systems of Thought: Holistic Versus Analytic Cognition, *Psychological Review*, *108* (2), 291-310.

Nordgren, L. F., van Harreveld, F. et van der Pligt, J. (2006). Ambivalence, Discomfort and Motivated Information Processing, *Journal of Experimental Social Psychology*, *42* (42), 252-258.

Norris, C. J., Larsen, J. T., Crawford, L. E. et Cacioppo, J. T. (2011). Better (or worse) for some than others: Individual differences in the positivity offset and negativity bias, *Journal of Research in Personality*, *45* (1), 100-111.

Oetting, E. R. (1967). The Effect of Forcing Response on Semantic Differential, *Journal of Educational and Psychological Measurement*, *27*, 699-702.

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions, *Journal of Marketing Research*, *17* (4), 460-469.

Oliver, R. L. (1997). *Satisfaction: A Behavioral Perspective on the Consumer*, New York, McGraw-Hill.

Olsen, S. O. (1999). Strength and Conflicting Valence in the Measurement of Food Attitudes and Preferences, *Food Quality and Preferences*, *10*, 483-494.

Olsen, S. O. (2002). Comparative Evaluation and the Relationship Between Quality, Satisfaction, and Repurchase Loyalty, *Journal of the Academy of Marketing Science*, *30* (3), 240-249.

Olsen, S. O., Wilcox, J. et Olsson, U. (2005). Consequences of Ambivalence on Satisfaction and Loyalty, *Psychology and Marketing*, *22* (3), 247-269.

Osgood, C. E. (1952). The Nature and Measurement of Meaning, *Psychological Bulletin*, *49* (3), 197.

Osgood, C. E., Suci, G. J. et Tannenbaum, P. H. (1957). *The measurement of meaning*, Urbana, University of Illinois press.

Otnes, C., Lowrey, T. M. et Shrum, L. J. (1997). Toward an Understanding of Consumer Ambivalence, *Journal of Consumer Research*, *24*, 80-93.

Parasuraman, A., Zeithaml, V. A. et Berry, L. L. (1988). SERVQUAL: a Multi-item Scale for Measuring Consumer Perceptions of the Service Quality, *Journal of Retailing, 64* (1), 12-40.

Parasuraman, A., Zeithaml, V. A. et Berry, L. L. (1994). Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research, *Journal of Marketing, 58* (January), 111-124.

Parguel, B. (2009). *L'influence de la communication promotionnelle sur le capital de la marque: le cas de la communication promotionnelle en prospectus en grande consommation*. Thèse de Doctorat, Université Paris-Dauphine.

Peterson, R. A. et Wilson, W. R. (1992). Measuring Customer Satisfaction: Fact and artifact, *Journal of the Academy of Marketing Science*, *20*, 61-71.

Petrocelli, J. V., Tormala, Z. L. et Rucker, D. D. (2007). Unpacking Attitude Certainty: Attitude Clarity and Attitude Correctness, *Journal of Personality and Social Psychology*, *92*, 30-41.

Petty, R. E. et Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*, New York, Springer.

Petty, R. E., Priester, J. R. et Wegener, D. T. (1994). Cognitive Processes in Attitude Change, *Handbook of social cognition*, *2*, 69-142.

Pinson, C. (1983). Pour une étude critique du différentiel sémantique, *Revue Française du Marketing, 95* (4), 3-27.

Preacher, K. J. et Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models., *Behavior Research Methods, 40*, 879-891.

Priester, J. R. et Petty, R. E. (1996). The Gradual Threshold Model of Ambivalence: Relating the Positive and Negative Bases of Attitudes to Subjective Ambivalence, *Journal of Personality and Social Psychology*, *71*, 431-449.

Priester, J. R. et Petty, R. E. (2001). Extending the bases of subjective attitudinal ambivalence: interpersonal and intrapersonal antecedents of evaluative tension, *Journal of Personality and Social Psychology*, *80*, 19-34.

Priester, J. R., Petty, R. E. et Park, K. (2007). Whence Univalent Ambivalence? From the Anticipation of Conflicting Reactions, *Journal of Consumer Research*, *34*, 11-21.

Ray, D. et Gotteland, D. (2005). Mesurer l'asymétrie des impacts des attributs sur la satisfaction : comparaison de la validité convergente de cinq méthodes, *Recherche et Applications en Marketing*, *20* (1), 1-19.

Reichheld, F. F. (2003). The One Number you Need to Grow, *Harvard Business Review*, *81* (12), 46-55.

Reisenzein, R. (1995). On Oatley and Johnson-Laird's Theory of Emotion and Hierarchical Structures in the Affective Lexicon, *Cognition and Emotion*, 383-416.

Richardson, L. (2000). Writting: A Model of Inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (2 ed.), CA: Thousand Oaks, Sage.

Roussel, P., Durrieu, F., Campoy, E. et El Akremi, A. (2002). *Méthodes d'équations structurelles : recherche et applications en gestion*, Paris, Economica.

Roussel, P. et Wacheux, F. (2005). *Management des ressources humaines : Méthodes de recherche en sciences humaines et sociales*, Bruxelles, De Boeck.

Rozin, P. et Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion, *Personality and Social Psychology Review*, *5* (4), 296-320.

Rucker, D. D. et Petty, R. E. (2004). When Resistance is Futile: Consequences of Failed Counterarguing for Attitude Certainty, *Journal of Personality and Social Psychology*, *86*, 219–235.

Russell, J. A. et Carroll, J. M. (1999). On the Bipolarity of Positive and Negative Affect, *Psychological Bulletin*, *125*, 3-30.

Sample, J. et Warland, R. (1973). Attitudes and the Prediction of Behaviour, *Social Forces*, *51* (Mach), 292-304.

Schneider, I. K., Eerland, A., van Harreveld, F., Rotteveel, M., van der Pligt, J., Van der Stoep, N. et Zwaan, R. (2013). One way and the other: The bidirectional relationship between ambivalence and body movement, *Psychological Science*, *24* (3), 319-325

Schoormans, J. P. L. et Robben, H. S. J. (1997). The effect of new package on product attention, categorization and evaluation, *Journal of Economic Psychology*, *18* (2), 271-287.

Scott, W. A. (1966). Brief report: Measures of cognitive structure, *Multivariate Behavioral Research*, *1* (3), 391-395.

Scott, W. A. (1968). Attitude measurement. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 2, pp. 204-273), MA: Addison-Wesley, Reading.

Seiders, K., Voss, G. B., Grewal, D. et Godfrey, A. L. (2005). Do Satisfied Customers Buy More? Examining Moderating Influences in a Retailing Context, *Journal of Marketing*, *69* (4), 26-43.

Sengupta, J. et Johar, G. V. (2002). The Effects of Inconsistent Attribute Information on the Predictive Value of Product Attitudes: Toward a Resolution of Opposing Perspectives, *Journal of Consumer Research*, *29*, 39-56.

Sincoff, J. B. (1990). The Psychological Characteristics of Ambivalent People, *Clinical Psychology Review*, *10* (1), 43-67.

Sparks, P., Conner, M., James, R., Shepherd, R. et Povey, R. (2001). Ambivalence about Health Related Behaviours: an Exploration in the Domain of Food Choice, *British Journal of Health Psychology*, *6*, 53-68.

Srull, T. K. et Wyer, R. S. (1989). Person memory and judgment, *Psychological Review*, *96* (1), 58.

Stragner, R. et Osgood, C. E. (1946). Impact of War on Nationalistic Frame of Reference: I. Changes in General Approval and Qualitative Patterning of Certain Stereotypes, *Journal of Social Psychology*, *24*, 187-215.

Suchman, E. (1950). The intensity component in attitude and opinion research. In S. Stoer & e. al. (Eds.), *Measurement and prediction*, Princeton, Princeton University Press.

Szymanski, D. M. et Henard, D. H. (2001). Customer Satisfaction: A Meta-Analysis of the Empirical Evidence, *Journal of the Academy of Marketing Science*, *29* (1), 16-35.

Taylor, S. A. (1997). Assessing Regression-Based Importance Weights for Quality Perceptions and Satisfaction Judgments in the Presence of Higher Order and/or Interaction Effects, *Journal of Retailing*, *73* (Spring), 135-159.

Thompson, M. M., Zanna, M. P. et Griffin, D. W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 361-386), Mahwah, NJ, Lawrence Erlbaum.

Thornton, J. R. (2011). Ambivalent or Indifferent? Examining the Validity of an Objective Measure of Partisan Ambivalence, *Political Psychology*, *32* (5), 863-884.

Ulin, P., Waszak, C. et Pfannenschmidt, S. (1996). *Integrating qualitative and quantitative research*. Paper presented at the Familiy Health International's Women's Studies Project Technical Advisory Group, Raleigh: NC.

van Harreveld, F., Schneider, I. K., Nohlen, H. U. et van der Pligt, J. (2012). The dynamics of ambivalence: Evaluative conflict in attitudes and decision making. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 267-284), New York, NY, The Guilford Press.

van Harreveld, F., van der Pligt, J. et de Liver, Y. N. (2009). The Agony of Ambivalence and Ways to Resolve it: Introducing the MAID Model, *Personality and Social Psychology Review*, *13*, 45-61.

van Harreveld, F., van der Pligt, J., de Vries, N. K., Wenneker, C. et Verhue, D. (2004). Ambivalence and Information Integration in Attitudinal Judgment, *British Journal of Social Psychology*, *43*, 431-447.

Velicer, W. F. et Fava, J. L. (1998). Effects of Variable and Subject Sampling on Factor Pattern Recovery, *Psychological Methods*, *3* (2), 231-251.

Westbrook, R. A. (1980). A Rating Scale for Measuring Product / Service Satisfaction, *Journal of Marketing*, *44* (Fall), 68-72.

Wright, P. (1973). Use of Consumer Judgement Models in Promotion Planning, *Journal of Marketing*, *37* (October), 27-33.

Yoo, S. (2010). Two Types of Neutrality: Ambivalence versus Indifference and Political Participation, *The Journal of Politics, 72* (1), 163-177.

Zaaraoui-Sabri, O. (2007). *Le tabou en communication publicitaire : étude des réactions affectives, cognitives et conatives du récepteur exposé à des stimuli tabou mortifères et sexuels.* Thèse de doctorat, Université Paris-Dauphine, Paris.

Zaichkowsky, J. L. (1985). Measuring the Involvement Construct, *Journal of Consumer Research*, *12*, 341-352.

Zaltman, G. (1997). Rethinking market research: Putting people back in, *Journal of Marketing Research*, 424-437.

Zanna, M. P. et Rempel, J. K. (1988). Attitudes: a new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge*, New York, Cambridge University Press.

Zeithaml, V. A., Berry, L. L. et Parasuraman, A. (1996). The Behavioral Consequences of Service Quality, *Journal of Marketing*, *60* (April), 31-46.

Zemborain, M. R. et Johar, G. V. (2007). Attitudinal ambivalence and openness to persuasion: A framework for interpersonal influence, *Journal of Consumer Research*, *33* (4), 506-514.

Zhao, X. et Cai, X. (2008). The Role of Ambivalence in College Nonsmokers' Information Seeking and Information Processing, *Communication Research*, *35*, 298-318.

Zhao, X., Lynch, J. G. et Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about Mediation Analysis, *Journal of Consumer Research*, *37* (2), 197-206.

– Annexes –

Annexe 1 : Questionnaire de l'étude 1

Page 1

Bonjour,

Vous allez participer à une étude portant sur les processus d' évaluation des individus.

Lisez attentivement les consignes qui vous seront données, svp.

Il n'y a pas de bonnes ni de mauvaises réponses, c'est votre OPINION PERSONNELLE et SUBJECTIVE qui nous intéresse.

DAUPHINE
UNIVERSITÉ PARIS

Cliquez sur le bouton "suivant" pour commencer, svp. Suivant >

Page 2

Imaginez que vous prévoyiez de faire un week-end en amoureux à La Rochelle cet été.

En vue de l'organisation de ce séjour qui comprendra une nuit d'hôtel et un repas au restaurant, vous vous rendez sur booking.com et vous consultez quelques avis de clients sur des hôtels et restaurants de la région.

Pouvez-vous nous donner l'opinion que vous vous faites de ces services suite à la lecture des commentaires ?

Suivant >

Page 3

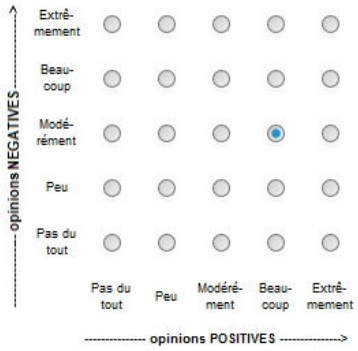
- *Condition ESG*

Dans cette étude nous allons vous demander d'utiliser une grille comme celle qui est présentée ci-dessous pour évaluer des services à l'aide d'avis de consommateurs.

Sur cette grille :
- l'axe horizontal correspond à la force de vos opinions POSITIVES
- l'axe vertical correspond à la force de vos opinions NEGATIVES

Vous pouvez ainsi cliquer sur le point qui décrit le mieux à votre opinion générale (positive et négative) suite à la lecture des avis de consommateurs.

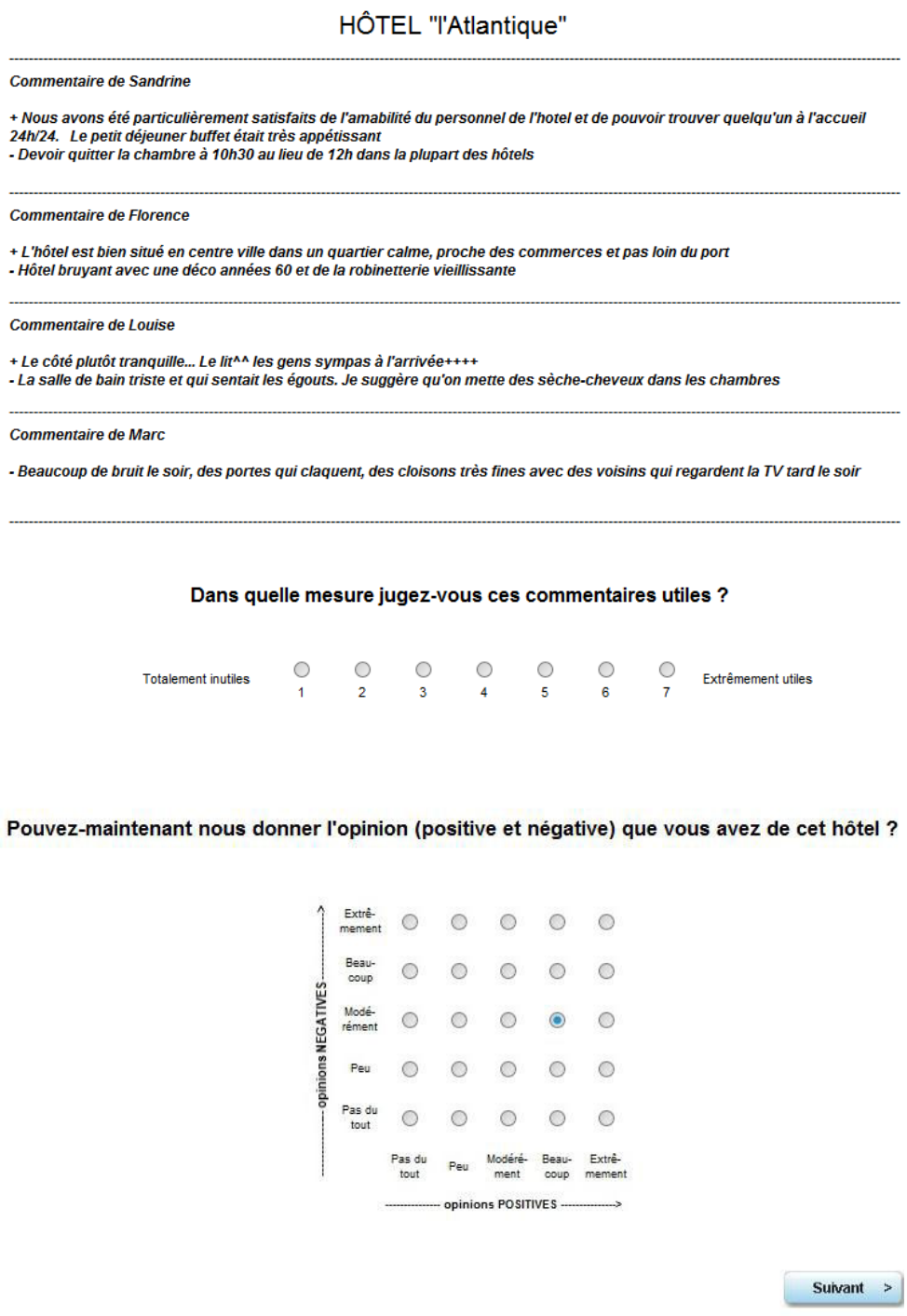
EXEMPLE :
Si vous jugez que les commentaires que vous avez lus pour un hôtel ou un restaurant sont dans l' ensemble très positifs, mais qu'ils soulignent tout de même un certain nombre de points faibles, vous pourrez cliquer sur le point marqué en bleu sur la grille ci-contre.



Suivant >

Page 4

- *Condition ESG*



Page 5

- *Condition ESG*

RESTAURANT « Le Nautilus »

Commentaire de Martine

De la bonne cuisine qui n'a rien de très original. Des plats principaux à la carte à 20€, mais dont les portions et les garnitures sont copieuses (ça rattrape). Service agréable. Beaucoup de monde le jour où nous y sommes allés.

Commentaire de Romane

Présentation des mets au top, raffinement et service également. Grande exigence au niveau des produits proposés (vin ++). Cadre sympa sans chichi. A recommander en pensant à réserver avant.

Commentaire de Thomas

Cuisine extra. Personnel très gentil. Nous nous sommes régalés. Petite ardoise de jambon avec apéritif . Le carpaccio de boeuf était super. Quant au tiramisu, nous en garderons un excellent souvenir. N'hesitez pas à y aller, vous ne serez pas déçus.

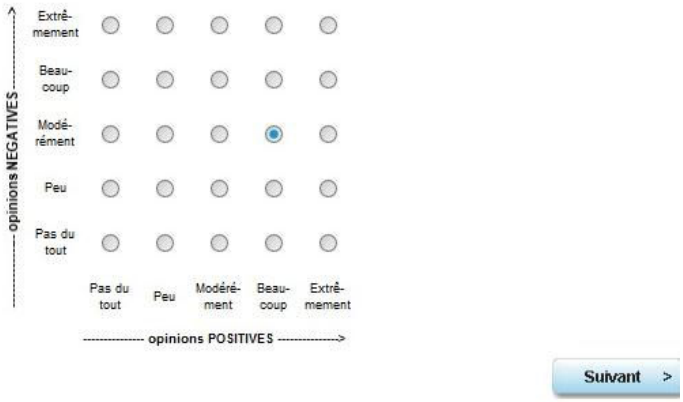
Commentaire de Christine

Service dynamique et très bien orchestré. Le choix est cornélien, tout donne envie. Nous avons opté en entrée pour le tartare de saumon et mousse d'avocat au wasabi, et la salade "fol dingue" aux gambas et cecina. C'était un véritable régal !

Dans quelle mesure jugez-vous ces commentaires utiles ?

Totalement inutiles ○1 ○2 ○3 ○4 ○5 ○6 ○7 Extrêmement utiles

Pouvez-vous maintenant nous donner l'opinion (positive et négative) que vous avez de ce restaurant ?



Suivant >

Page 6

- *Condition ESG*

RESTAURANT « Le Grand Large »

Commentaire de Juliette

Un accueil et un service exécrable, on se serait cru à la cantine! Dès 21h le soir il n'est plus possible de boire une bière. La nourriture était clairement trop cuite et réchauffée (les Saint jacques auraient pu rebondir comme une balle de caoutchouc) Le final a été au moment du café lorsque les desserts à peine servi à 22h30, il a fallu commander les cafés car ils fermaient la machine...

Commentaire de David

Cuisine à base de produits frais correcte. Service un peu lent.

Commentaire de Stéphanie

Personnel peu agréable et coquillages un peu secs à force d'être réchauffés. Pas beaucoup de choix concernant les desserts. Nous avons été déçus.

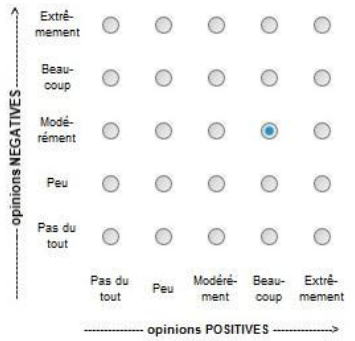
Commentaires d'Alexandra

Les plats sont joliment présentés mais sans saveur: terrine de poissons insipide, côte de veau sèche cachée par une sauce. Le steak tartare était le plus mauvais aussi loin que je me rappelle. Le service est lent. Bref, à éviter...

Dans quelle mesure jugez-vous ces commentaires utiles ?

Totalement inutiles ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 Extrêmement utiles

Pouvez-vous maintenant nous donner l'opinion (positive et négative) que vous avez de ce restaurant ?



Suivant >

Page 3

- Condition 2CU

HÔTEL "l'Atlantique"

Commentaire de Sandrine

+ Nous avons été particulièrement satisfaits de l'amabilité du personnel de l'hotel et de pouvoir trouver quelqu'un à l'accueil 24h/24. Le petit déjeuner buffet était très appétissant
- Devoir quitter la chambre à 10h30 au lieu de 12h dans la plupart des hôtels

---

Commentaire de Florence

+ L'hôtel est bien situé en centre ville dans un quartier calme, proche des commerces et pas loin du port
- Hôtel bruyant avec une déco années 60 et de la robinetterie vieillissante

---

Commentaire de Louise

+ Le côté plutôt tranquille... Le lit^^ les gens sympas à l'arrivée++++
- La salle de bain triste et qui sentait les égouts. Je suggère qu'on mette des sèche-cheveux dans les chambres

---

Commentaire de Marc

- Beaucoup de bruit le soir, des portes qui claquent, des cloisons très fines avec des voisins qui regardent la TV tard le soir

---

Dans quelle mesure jugez-vous ces commentaires utiles ?

Totalement inutiles ○1 ○2 ○3 ○4 ○5 ○6 ○7 Extrêmement utiles

Pouvez-vous maintenant nous donner l'opinion que vous avez de cet hôtel ?

Dans quelle mesure avez-vous une opinion positive de cet hôtel (en vous concentrant uniquement sur les aspects postifs de l'hôtel et en ignorant les aspects négatifs) ?

Pas du tout positive ○1 ○2 ○3 ○4 ○5 Extrêmement positive

Dans quelle mesure avez-vous une opinion négative de cet hôtel (en vous concentrant uniquement sur les aspects négatifs de l'hôtel et en ignorant les aspects positifs) ?

Pas du tout négative ○1 ○2 ○3 ○4 ○5 Extrêmement négative

Suivant >

Page 4

- Condition 2CU

RESTAURANT « Le Nautilus »

Commentaire de Martine

De la bonne cuisine qui n'a rien de très original. Des plats principaux à la carte à 20€, mais dont les portions et les garnitures sont copieuses (ça rattrape). Service agréable. Beaucoup de monde le jour où nous y sommes allés.

Commentaire de Romane

Présentation des mets au top, raffinement et service également. Grande exigence au niveau des produits proposés (vin ++). Cadre sympa sans chichi. A recommander en pensant à réserver avant.

Commentaire de Thomas

Cuisine extra. Personnel très gentil. Nous nous sommes régalés. Petite ardoise de jambon avec apéritif . Le carpaccio de boeuf était super. Quant au tiramisu, nous en garderons un excellent souvenir. N'hesitez pas à y aller, vous ne serez pas déçus.

Commentaire de Christine

Service dynamique et très bien orchestré. Le choix est cornélien, tout donne envie. Nous avons opté en entrée pour le tartare de saumon et mousse d'avocat au wasabi, et la salade "fol dingue" aux gambas et cecina. C'était un véritable régal !

Dans quelle mesure jugez-vous ces commentaires utiles ?

Totalement inutiles	○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	Extrêmement utiles

Pouvez-vous maintenant nous donner l'opinion que vous avez de ce restaurant ?

Dans quelle mesure avez-vous une opinion positive de ce restaurant (en vous concentrant uniquement sur les aspects postifs du restaurant et en ignorant les aspects négatifs) ?

Pas du tout positive	○ 1	○ 2	○ 3	○ 4	○ 5	Extrêmement positive

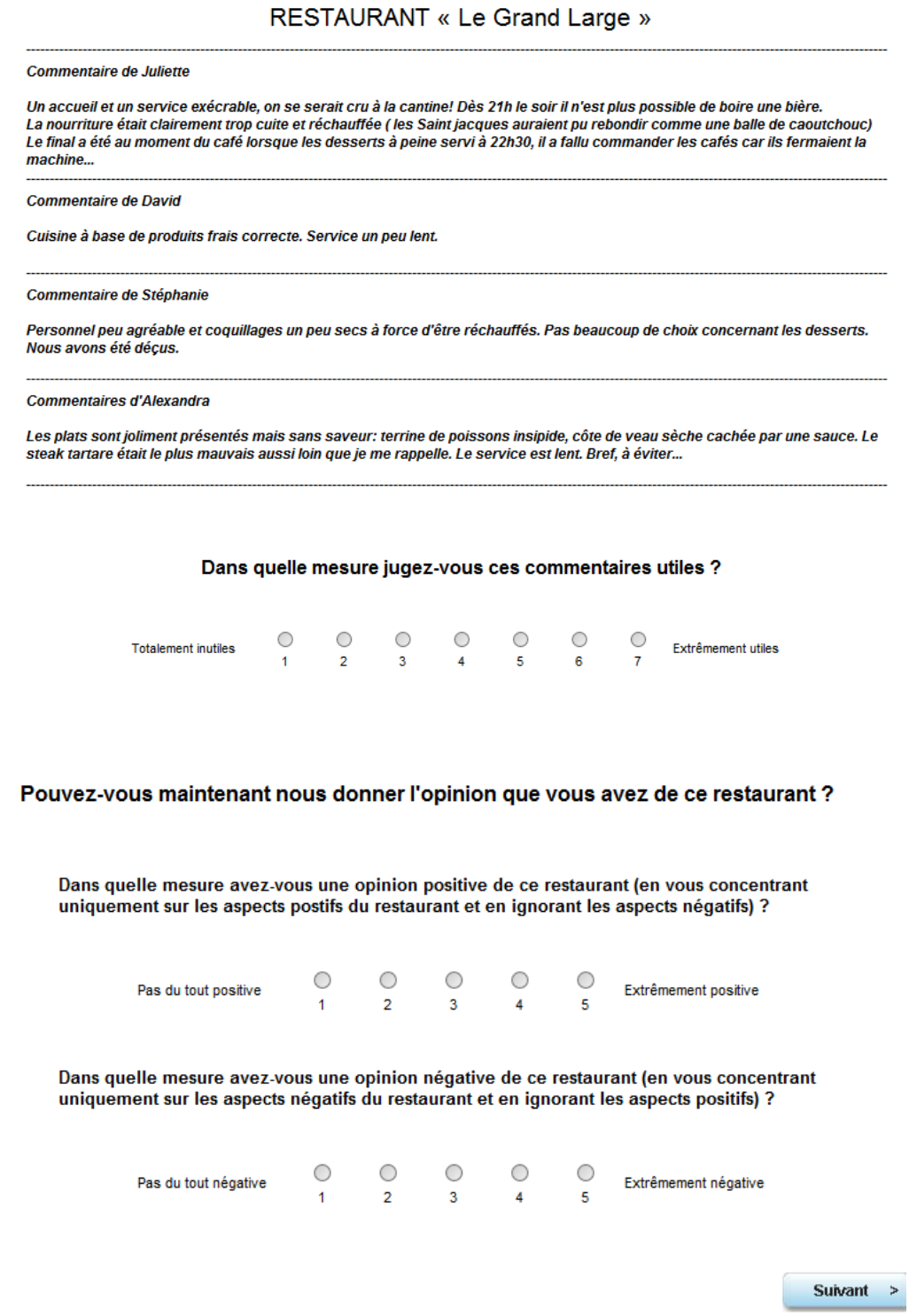
Dans quelle mesure avez-vous une opinion négative de ce restaurant (en vous concentrant uniquement sur les aspects négatifs du restaurant et en ignorant les aspects positifs) ?

Pas du tout négative	○ 1	○ 2	○ 3	○ 4	○ 5	Extrêmement négative

Suivant >

Page 5

- Condition 2CU



RESTAURANT « Le Grand Large »

Commentaire de Juliette

Un accueil et un service exécrable, on se serait cru à la cantine! Dès 21h le soir il n'est plus possible de boire une bière.
La nourriture était clairement trop cuite et réchauffée (les Saint jacques auraient pu rebondir comme une balle de caoutchouc)
Le final a été au moment du café lorsque les desserts à peine servi à 22h30, il a fallu commander les cafés car ils fermaient la machine...

Commentaire de David

Cuisine à base de produits frais correcte. Service un peu lent.

Commentaire de Stéphanie

Personnel peu agréable et coquillages un peu secs à force d'être réchauffés. Pas beaucoup de choix concernant les desserts. Nous avons été déçus.

Commentaires d'Alexandra

Les plats sont joliment présentés mais sans saveur: terrine de poissons insipide, côte de veau sèche cachée par une sauce. Le steak tartare était le plus mauvais aussi loin que je me rappelle. Le service est lent. Bref, à éviter...

Dans quelle mesure jugez-vous ces commentaires utiles ?

Totalement inutiles ○1 ○2 ○3 ○4 ○5 ○6 ○7 Extrêmement utiles

Pouvez-vous maintenant nous donner l'opinion que vous avez de ce restaurant ?

Dans quelle mesure avez-vous une opinion positive de ce restaurant (en vous concentrant uniquement sur les aspects postifs du restaurant et en ignorant les aspects négatifs) ?

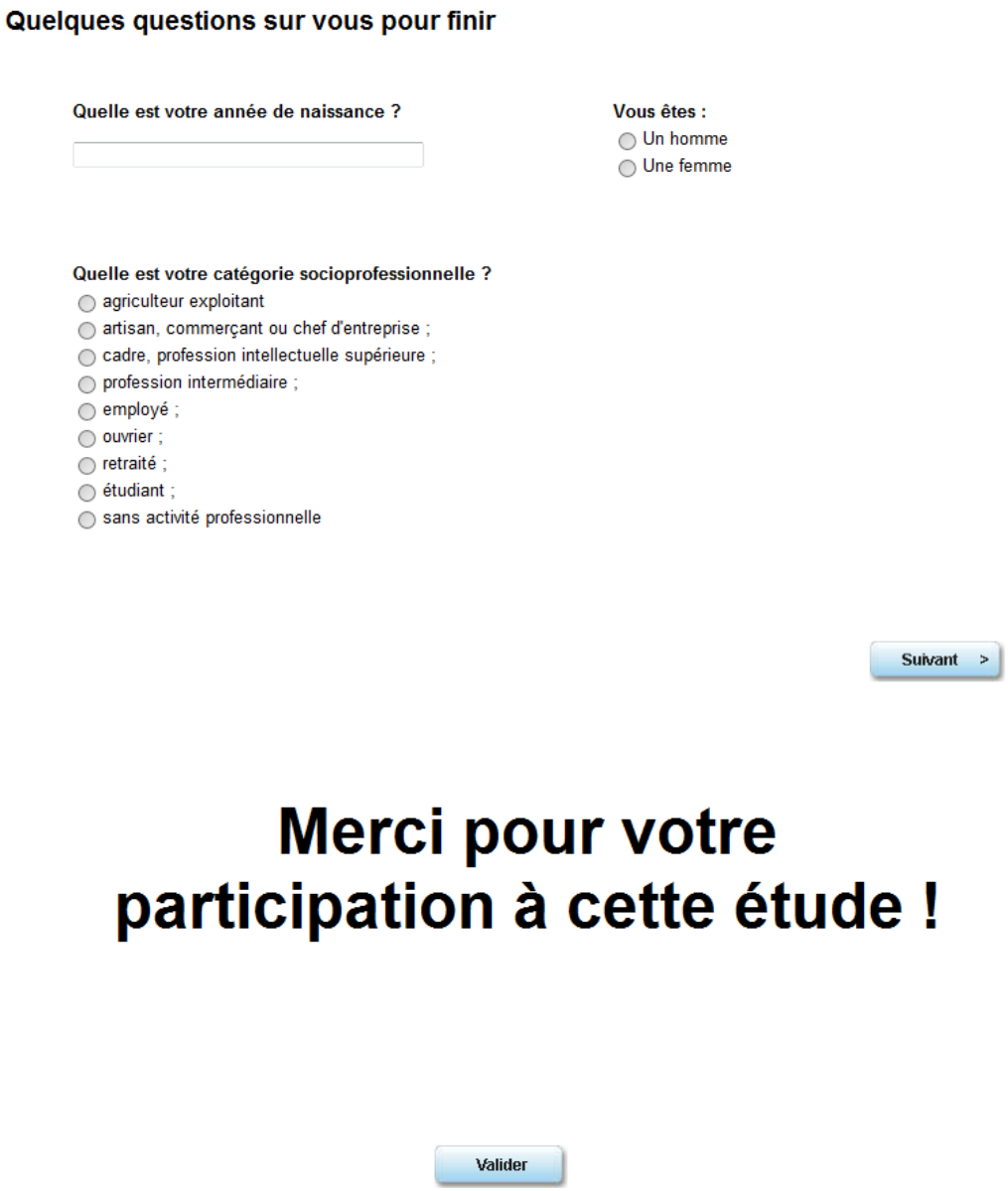
Pas du tout positive ○1 ○2 ○3 ○4 ○5 Extrêmement positive

Dans quelle mesure avez-vous une opinion négative de ce restaurant (en vous concentrant uniquement sur les aspects négatifs du restaurant et en ignorant les aspects positifs) ?

Pas du tout négative ○1 ○2 ○3 ○4 ○5 Extrêmement négative

Suivant >

La suite du questionnaire est identique pour les deux conditions expérimentales



Quelques questions sur vous pour finir

Quelle est votre année de naissance ?

Vous êtes :

- Un homme
- Une femme

Quelle est votre catégorie socioprofessionnelle ?

- agriculteur exploitant
- artisan, commerçant ou chef d'entreprise ;
- cadre, profession intellectuelle supérieure ;
- profession intermédiaire ;
- employé ;
- ouvrier ;
- retraité ;
- étudiant ;
- sans activité professionnelle

Suivant >

Merci pour votre participation à cette étude !

Valider

DAUPHINE UNIVERSITÉ PARIS

Annexe 2 : Questionnaire de l'étude 2

Page 1

Bonjour,

Merci d'avoir accepté de répondre à ce questionnaire.

Cette étude est menée par un centre de RECHERCHE UNIVERSITAIRE.
Elle n'a AUCUNE FINALITE COMMERCIALE. Nous nous intéressons simplement aux processus d'évaluation des individus.

Les données collectées sont totalement anonymes et ne seront diffusées que dans des REVUES SCIENTIFIQUES spécialisées.

Nous vous demandons simplement de répondre le plus sincèrement et spontanément possible.

Cela vous prendra environ 10 MINUTES.

Cliquez sur le bouton "suivant" pour commencer, svp.

Suivant >

Page 2

Dans cette étude, nous allons vous demander d'évaluer 2 services en utilisant une grille identique à celle que vous pouvez voir ci-dessous.

Cette grille permet d'évaluer PLUSIEURS ELEMENTS en même temps :
- l'AXE VERTICAL correspond à votre niveau d'INSATISFACTION ;
- l'AXE HORIZONTAL correspond à votre niveau de SATISFACTION.

Vous pouvez ainsi cliquer sur la pastille qui décrit le mieux à votre opinion générale



Par exemple, quand vous allez au restaurant :
- Vous pouvez à la fois être satisfait(e) par le service et la qualité des plats ;
- et un peu insatisfait(e) par la musique diffusée dans le restaurant.
(cf. exemple de la pastille cochée ci-contre)

Suivant >

Page 3

Quelques petits exemples pour vous entrainer… (1/3)

D'après vous, quel est le niveau de satisfaction / insatisfaction de la personne ayant laissé le commentaire suivant sur lafourchette.com ?

"1h d'attente - restau bondé , ultra bruyant , serveur détestable . La nourriture est correcte mais ça n'en vaut vraiment pas la peine si c'est pour se faire maltraiter pendant une heure et devoir partir immédiatement après avoir bu la dernière gorgée de son café !"



Suivant >

Page 4

Quelques petits exemples pour vous entrainer… (1/3)

Evidemment, il n'y a pas de bonne réponse dans l'absolu...

Mais on peut imaginer que la personne qui a écrit ce commentaire situerait son évaluation dans cette zone de la grille.

"1h d'attente - restau bondé , ultra bruyant , serveur détestable . La nourriture est correcte mais ça n'en vaut vraiment pas la peine si c'est pour se faire maltraiter pendant une heure et devoir partir immédiatement après avoir bu la dernière gorgée de son café !"



Suivant >

Page 5

Quelques petits exemples pour vous entrainer… (2/3)

D'après vous, quel est le niveau de satisfaction / insatisfaction de la personne ayant laissé le commentaire suivant sur lafourchette.com ?

"Très bon restaurant pour un prix plus qu'abordable. Salade vraiment bien garnie (gésiers, foie, oeuf au plat, cantal...). Ca fait plaisir d'avoir des produits préparés par le chef et pas un mélange de différentes boîtes ou produits surgelés !"



Suivant >

Page 6

Quelques petits exemples pour vous entrainer… (2/3)

Evidemment, il n'y a pas de bonne réponse dans l'absolu...

Mais on peut imaginer que la personne qui a écrit ce commentaire situerait son évaluation dans cette zone de la grille.

"Très bon restaurant pour un prix plus qu'abordable. Salade vraiment bien garnie (gésiers, foie, oeuf au plat, cantal...). Ca fait plaisir d'avoir des produits préparés par le chef et pas un mélange de différentes boîtes ou produits surgelés !"



Suivant >

Page 7

Quelques petits exemples pour vous entrainer… (3/3)

D'après vous, quel est le niveau de satisfaction / insatisfaction de la personne ayant laissé le commentaire suivant sur lafourchette.com ?

"Pas trop d'intimité dans ce restaurant surbondé. En revanche c'est convivial, on mange bien, et l'addition sait rester légère !
Petit doute tout de même sur les desserts qui sont écoeurants à mon goût…
Mieux vaut arriver à 19h pour ne pas attendre une table pendant 1h"



Suivant >

Page 8

Quelques petits exemples pour vous entrainer… (3/3)

Evidemment, il n'y a pas de bonne réponse dans l'absolu...

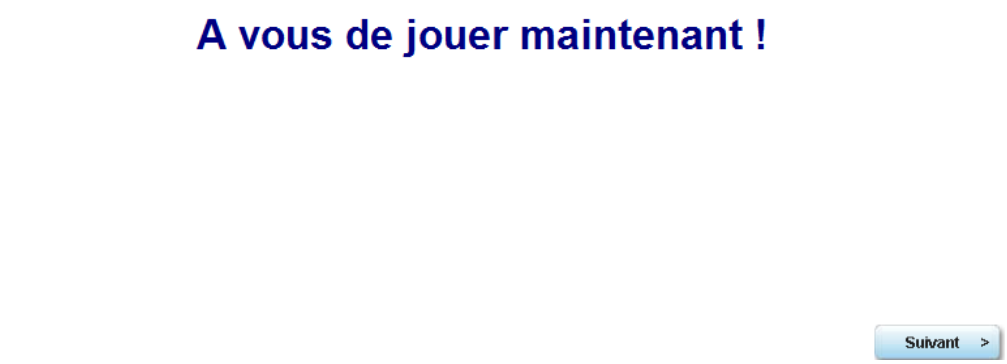
Mais on peut imaginer que la personne qui a écrit ce commentaire situerait son évaluation dans cette zone de la grille.

"Pas trop d'intimité dans ce restaurant surbondé. En revanche c'est convivial, on mange bien, et l'addition sait rester légère !
Petit doute tout de même sur les desserts qui sont écoeurants à mon goût…
Mieux vaut arriver à 19h pour ne pas attendre une table pendant 1h"



Suivant >

Page 9



Page 10



Page 11

Si vous deviez évaluer vos expériences dans les restaurants Mc Donald's...

Plus précisément, quel est votre niveau de satisfaction / insatisfaction vis-à-vis de vos expériences au McDonald's ?



Pensez-vous que cette réponse reflète parfaitement votre opinion ?

------> *Pas du tout* *Tout à fait*

Suivant >

Page 12

Si vous deviez évaluer vos expériences dans les restaurants Mc Donald's...

Pouvez-vous préciser un peu votre opinion concernant vos expériences au MC DONALD'S ?

Quels sont les éléments qui vous rendent INSATISFAIT quand vous allez au Mc Donald's ?

Quels sont les éléments qui vous rendent SATISFAIT quand vous allez au Mc Donald's ?

Suivant >

Page 13

Si vous deviez évaluer vos expériences dans les restaurants Mc Donald's...

Etes-vous d'accord avec les affirmations suivantes concernant vos expériences au MC DONALD'S ?

	Pas du tout d' accord						Tout à fait d' accord
J'ai des opinions contradictoires au sujet du Mc Donald's	○	○	○	○	○	○	○
J'ai un avis partagé vis-à-vis du Mc Donald's	○	○	○	○	○	○	○
Je suis indécis au sujet du Mc Donald's	○	○	○	○	○	○	○
J'ai des opinions positives et négatives au sujet du Mc Donald's	○	○	○	○	○	○	○

Suivant >

Page 14

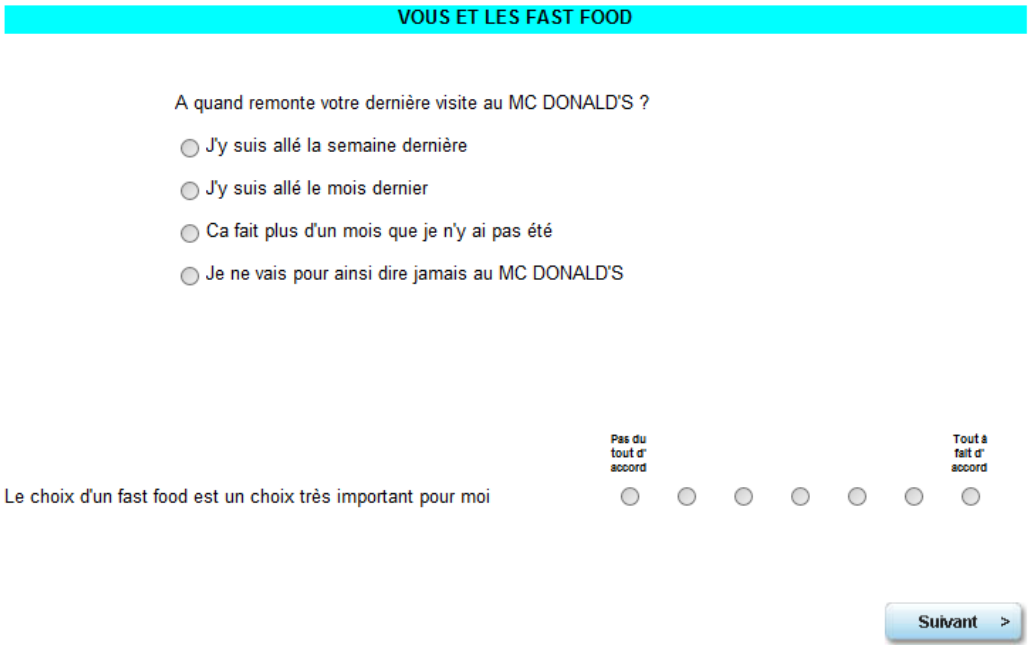
Comment comptez-vous vous comporter dans le futur vis-à-vis du Mc Donald's ?

	Pas du tout probable						Extrê-ment probable
J'ai l'intention de dire des choses positives aux gens sur le Mc Donald's	○	○	○	○	○	○	○
Je désire retourner au Mc Donald's dans le futur	○	○	○	○	○	○	○
Je vais encourager mes amis à aller au Mc Donald's	○	○	○	○	○	○	○
J'ai l'intention de parler du Mc Donald's en termes peu flatteurs	○	○	○	○	○	○	○

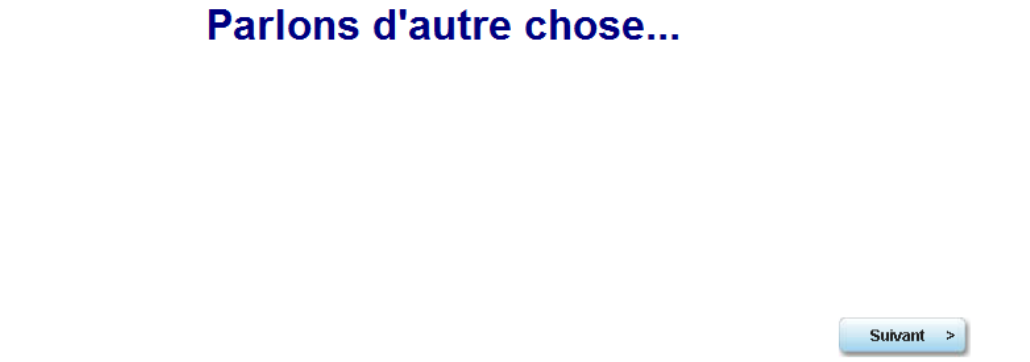
	Pas du tout probable						Extrê-ment probable
J'ai l'intention d'aller au Mc Donald's dans le futur	○	○	○	○	○	○	○
Je vais recommander le Mc Donald's aux personnes qui me demandent conseil	○	○	○	○	○	○	○
Je vais continuer à être client du Mc Donald's	○	○	○	○	○	○	○
Je compte dire principalement des choses négatives aux autres sur le Mc Donald's	○	○	○	○	○	○	○

Suivant >

Page 15



Page 16



Page 17

Si vous deviez évaluer votre dernière expérience chez le COIFFEUR...

Quel est votre niveau de satisfaction vis-à-vis de cette dernière expérience chez le coiffeur ?

Extrême INSATISFACTION ○ ○ ○ ○ ○ ○ ○ ○ ○ Extrême SATISFACTION

Pensez-vous que cette réponse reflète parfaitement votre opinion ?

Pas du tout ○ ○ ○ ○ ○ *Tout à fait*

Suivant >

Page 18

Si vous deviez évaluer votre dernière expérience chez le COIFFEUR...

Plus précisément, quel est votre niveau de satisfaction / insatisfaction vis-à-vis de cette expérience chez le coiffeur ?



Pensez-vous que cette réponse reflète parfaitement votre opinion ?

------> *Pas du tout* ○ ○ ○ ○ ○ *Tout à fait*

Suivant >

Page 19

Si vous deviez évaluer votre dernière expérience chez le COIFFEUR...

Pouvez-vous préciser un peu votre opinion concernant votre dernière expérience chez le coiffeur ?

Quels sont les éléments qui ont contribué à votre INSATISFACTION la dernière fois que vous êtes allé chez le coiffeur ?

Quels sont les éléments qui ont contribué à votre SATISFACTION la dernière fois que vous êtes allé chez le coiffeur ?

Suivant >

Page 20

Si vous deviez évaluer votre dernière expérience chez le COIFFEUR...

Etes-vous d'accord avec les affirmations suivantes concernant votre dernière expérience chez le coiffeur ?

	Pas du tout d' accord						Tout à fait d' accord
J'ai des opinions contradictoires au sujet de ma dernière expérience chez le coiffeur	○	○	○	○	○	○	○
J'ai un avis partagé vis-à-vis de ma dernière expérience chez le coiffeur	○	○	○	○	○	○	○
Je suis indécis(e) au sujet de ma dernière expérience chez le coiffeur	○	○	○	○	○	○	○
J'ai des opinions positives et négatives au sujet de ma dernière expérience chez le coiffeur	○	○	○	○	○	○	○

Suivant >

Page 21

Comment comptez-vous vous comporter dans le futur vis-à-vis du coiffeur chez qui vous êtes allés la dernière fois ?

	Pas du tout probable						Extrê-mement probable
J'ai l'intention de dire des choses positives aux gens sur ce coiffeur	○	○	○	○	○	○	○
Je désire retourner chez ce coiffeur dans le futur	○	○	○	○	○	○	○
Je vais encourager mes amis à aller chez ce coiffeur	○	○	○	○	○	○	○
J'ai l'intention de parler de ce coiffeur en termes peu flatteurs	○	○	○	○	○	○	○

	Pas du tout probable						Extrê-mement probable
J'ai l'intention d'aller chez ce coiffeur dans le futur	○	○	○	○	○	○	○
Je vais recommander ce coiffeur aux personnes qui me demandent conseil	○	○	○	○	○	○	○
Je vais continuer à être client de ce coiffeur	○	○	○	○	○	○	○
Je compte dire principalement des choses négatives aux autres sur ce coiffeur	○	○	○	○	○	○	○

Suivant >

Page 22

VOUS ET LE COIFFEUR

A quand remonte la dernière fois que vous êtes allés chez le coiffeur ?

- ○ J'y suis allé le mois dernier
- ○ J'y suis allé il y a moins de 6 mois
- ○ Ca fait plus de 6 mois que je n'y ai pas été
- ○ Je ne vais pour ainsi dire jamais chez le coiffeur

	Pas du tout d' accord						Tout à fait d' accord
Le choix de mon coiffeur est un choix très important pour moi	○	○	○	○	○	○	○

Suivant >

Page 23

QUELQUES AUTRES QUESTIONS SUR VOUS POUR TERMINER

Quelle est votre année de naissance ?

Vous êtes :

Quel est votre dipôme le plus élevé ?

- Baccalauréat général ou professionnel
- BTS / DUT
- Licence
- Master 1
- Master 2 ou diplôme plus élevé

MERCI !!!

Valider

Annexe 3 : Questionnaire de l'étude 3

Page 1

Bonjour,

Vous allez participer une étude portant sur l'évaluation de votre banque.

Lisez chaque question attentivement et prenez le temps de bien NUANCER vos réponses.

Les réponses sont strictement anonymes et confidentielles. Il n'y a pas de bonnes ou de mauvaises réponses, c'est toujours votre OPINION PERSONNELLE et SUBJECTIVE qui nous intéresse.

Merci d'avance pour votre participation !

Cliquez sur le bouton "suivant" pour commencer, svp.

Suivant >

Page 2

QUELQUES QUESTIONS SUR VOUS POUR COMMENCER

Quel âge avez-vous ?

- ◯ Entre 18 et 29 ans
- ◯ Entre 30 et 40 ans
- ◯ Entre 41 et 50 ans
- ◯ Plus de 50 ans

Vous êtes :

- ◯ Un homme
- ◯ Une femme

Suivant >

Page 3

Des questions sur votre BANQUE PRINCIPALE (celle avec laquelle vous réalisez le plus de transactions) vont suivre.

Prenez quelques instants pour penser à votre expérience avec celle-ci.

Cliquez sur le bouton "suivant" quand vous serez prêt à répondre à des questions sur ce sujet.

Suivant >

Page 4

VOUS ET VOTRE BANQUE...

Depuis combien de temps êtes-vous client de cette banque ?

- ◯ Depuis 0 à 3 mois
- ◯ Depuis 4 à 6 mois
- ◯ Depuis 7 mois à 1 an
- ◯ Depuis 1 à 2 ans
- ◯ Depuis 3 à 5 ans
- ◯ Depuis + de 5 ans

Comment qualifieriez-vous votre niveau d'implication vis-à-vis de votre banque ?

C'est sans importance pour moi	◯ ◯ ◯ ◯ ◯ ◯ ◯	C'est très important pour moi
Cela ne compte pas pour moi	◯ ◯ ◯ ◯ ◯ ◯ ◯	Cela compte beaucoup pour moi
Cela ne me concerne pas	◯ ◯ ◯ ◯ ◯ ◯ ◯	Cela me concerne beaucoup
C'est négligeable	◯ ◯ ◯ ◯ ◯ ◯ ◯	C'est considérable
C'est inutile	◯ ◯ ◯ ◯ ◯ ◯ ◯	C'est utile

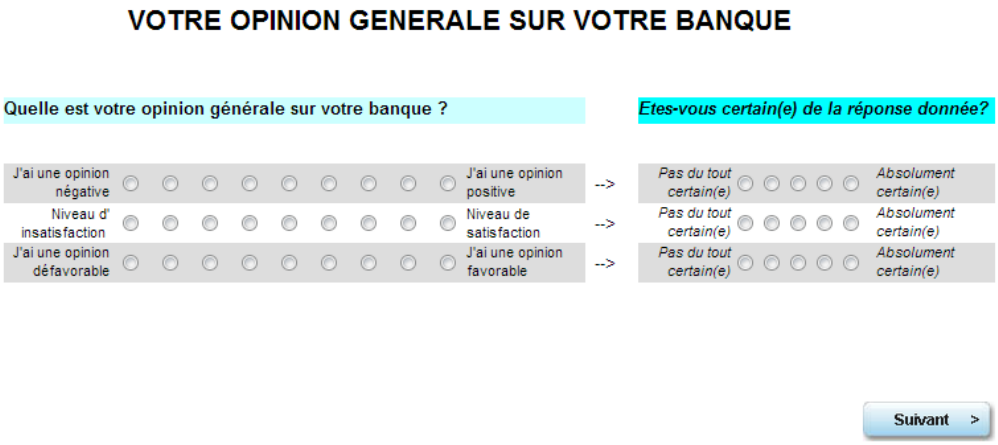
Comment qualifieriez-vous votre connaissance du secteur bancaire ?

J'ai très peu de connaissances sur le secteur bancaire	◯ ◯ ◯ ◯ ◯ ◯ ◯	J'ai beaucoup de connaissances sur le secteur bancaire
Je suis inexpérimenté(e) au sujet des banques	◯ ◯ ◯ ◯ ◯ ◯ ◯	Je suis expérimenté(e) au sujet des banques
Je ne suis pas informé(e) sur les produits bancaires	◯ ◯ ◯ ◯ ◯ ◯ ◯	Je suis informé(e) sur les produits bancaires
Je suis un(e) client(e) novice	◯ ◯ ◯ ◯ ◯ ◯ ◯	Je suis un(e) client(e) expert(e)

Suivant >

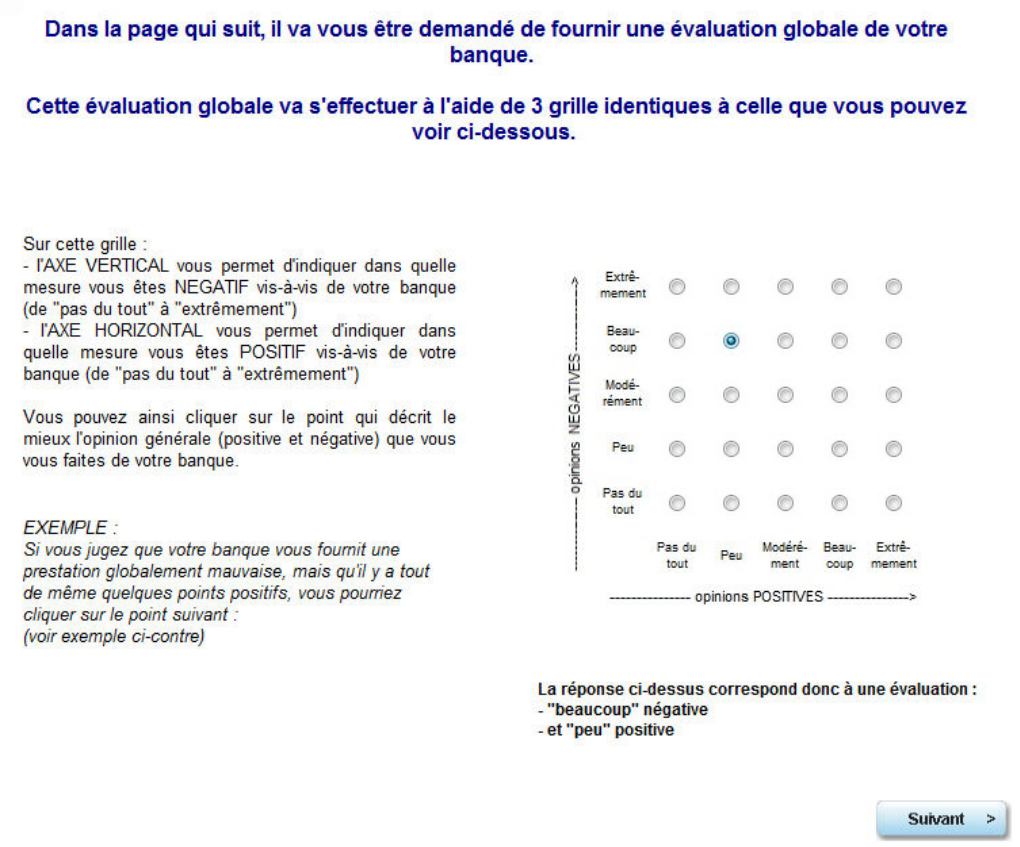
Page 5

Condition DS



Page 5

Condition ESG



Page 6 (suite *condition ESG*)

QUELLE EST VOTRE OPINION GENERALE SUR VOTRE BANQUE ?

Opinions négatives / opinions positives :

Pensez-vous que cette réponse reflète parfaitement votre opinion ?



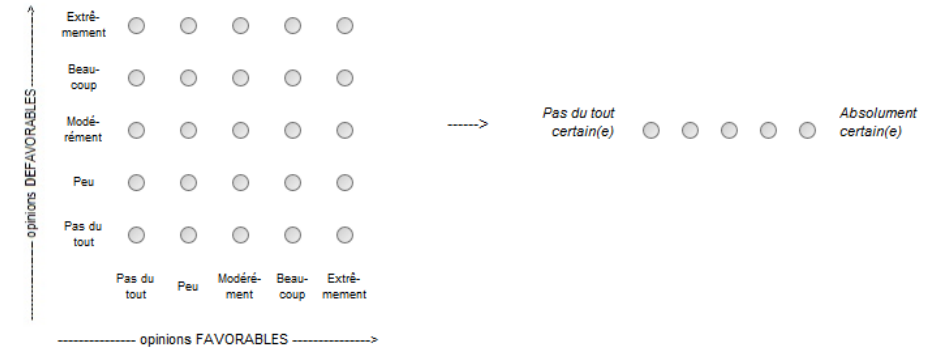
Insatisfaction / satisfaction :

Pensez-vous que cette réponse reflète parfaitement votre opinion ?



Opinions défavorables / opinions favorables :

Pensez-vous que cette réponse reflète parfaitement votre opinion ?



Suivant >

Page 7

(le questionnaire est identique pour les deux conditions à partir d'ici)

PLUS PRECISEMENT, VOTRE OPINION SUR VOTRE BANQUE...

Etes-vous d'accord avec les affirmations suivantes concernant votre banque ?

Concernant l'environnement des agences de cette banque	Pas du tout d' accord						Tout à fait d' accord
L'extérieur de votre agence bancaire est attrayant	○	○	○	○	○	○	○
L'intérieur de votre agence bancaire est attrayant	○	○	○	○	○	○	○
Les employés de cette banque ont une apparence propre et soignée	○	○	○	○	○	○	○
L'intérieur de votre agence bancaire est spacieux	○	○	○	○	○	○	○

Concernant les relations avec les employés de cette banque	Pas du tout d' accord						Tout à fait d' accord
Les employés de cette banque ont les connaissances nécessaires pour répondre aux problèmes	○	○	○	○	○	○	○
Les employés de cette banque sont polis avec les clients	○	○	○	○	○	○	○
Les employés de cette banques sont expérimentés	○	○	○	○	○	○	○
Les employés de cette banques inspirent la confiance aux clients	○	○	○	○	○	○	○
Les employés de cette banque sont à l'écoute des clients	○	○	○	○	○	○	○
Les employés de cette banque servent les clients de manière satisfaisante	○	○	○	○	○	○	○
Les employés de cette banque sont chaleureux avec leurs clients	○	○	○	○	○	○	○

Concernant les compétences techniques des employés de cette banque	Pas du tout d' accord						Tout à fait d' accord
Cette banque ne laisse pas ses clients faire la queue pendant un temps interminable	○	○	○	○	○	○	○
Les employés de cette banque réalisent les opérations demandées en temps voulu	○	○	○	○	○	○	○
Les employés de cette banque aident systématiquement les clients	○	○	○	○	○	○	○
Les employés de cette banque fournissent une attention personnalisée aux clients	○	○	○	○	○	○	○
Les employés de cette banque sont disposés à résoudre les problèmes des clients	○	○	○	○	○	○	○

Concernant la fiabilité de cette banque	Pas du tout d' accord						Tout à fait d' accord
Les employés de cette banque ne font pas d'erreur	○	○	○	○	○	○	○
Les employés de cette banque respectent la confidentialité des transactions	○	○	○	○	○	○	○
Les employés de cette banque fournissent des informations précises aux clients	○	○	○	○	○	○	○
Cette banque informe les clients de leurs opérations financières avec précision	○	○	○	○	○	○	○

Suivant >

Page 8

VOS INTENTIONS VIS-A-VIS DE VOTRE BANQUE...

Comment comptez-vous vous comporter vis-à-vis de votre banque dans les années à venir ?

	Pas du tout probable						Extrê-mement probable
J'ai l'intention de dire des choses positives aux gens sur cette banque dans les années à venir	○	○	○	○	○	○	○
Je vais encourager mes amis et mes proches à faire appel à cette banque	○	○	○	○	○	○	○
Je vais recommander cette banque aux personnes qui me demandent conseil	○	○	○	○	○	○	○

	Pas du tout probable						Extrê-mement probable
J'ai l'intention d'utiliser les services de cette banque dans les années à venir	○	○	○	○	○	○	○
Je souhaiterai utiliser les services de cette banque dans les années à venir	○	○	○	○	○	○	○
Je vais utiliser les services de cette banque dans les années à venir	○	○	○	○	○	○	○

Suivant >

Page 9

PLUS PRECISEMENT, VOTRE OPINION GENERALE SUR VOTRE BANQUE...

Etes-vous d'accord avec les affirmations suivantes concernant votre banque ?

	Pas du tout d' accord						Tout à fait d' accord
J'ai des opinions contradictoires au sujet de ma banque	○	○	○	○	○	○	○
J'ai un avis partagé vis-à-vis de ma banque	○	○	○	○	○	○	○
J'ai des opinions et des sentiments contradictoires au sujet de ma banque	○	○	○	○	○	○	○
Je suis indécis au sujet de ma banque	○	○	○	○	○	○	○
J'ai des opinions positives et négatives au sujet de ma banque	○	○	○	○	○	○	○

Suivant >

Page 10

QUELQUES AUTRES QUESTIONS SUR VOUS POUR TERMINER

Quelle est votre année de naissance ?

Merci d'indiquer votre situation familiale :

- Seul(e) sans enfant
- Seul(e) avec enfant(s)
- En couple sans enfant
- En couple avec enfant(s)

Quel est votre dipôme le plus élevé ?

- Aucun diplôme
- Brevet des collèges
- CAP / BEP
- Baccalauréat général ou professionnel
- BTS / DUT
- Licence
- Maîtrise / Diplôme d'école de commerce ou d'ingénieur
- Master / DEA / DESS
- Doctorat

Valider

Annexe 4 : Guide d'entretien de l'étude 4
(Première partie : qu'en pensent les chercheurs ?)

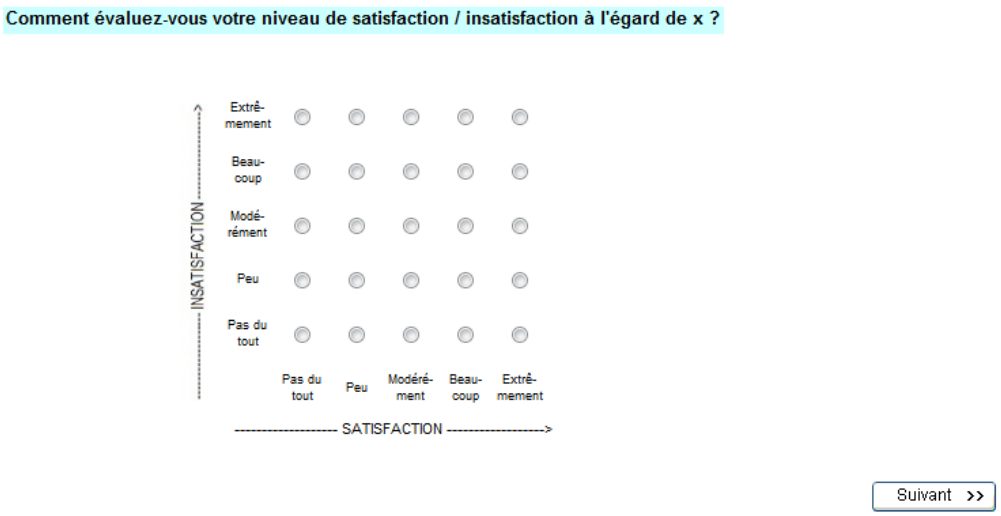
Introduction de l'entretien

Durant cet entretien, il va vous être demandé de répondre à différentes questions concernant votre niveau de satisfaction ou d'insatisfaction vis-à-vis de certains services.

Dans ma thèse, j'utilise un nouvel outil de mesure de la satisfaction. Cet entretien vise à comprendre comment les gens réagissent lorsqu'ils se trouvent face à cet outil de mesure qu'ils ne connaissent pas. J'aimerais comprendre le « cheminement mental » que les gens effectuent pour produire une réponse avec ce nouvel outil de mesure.

Donc ce qui m'intéresse, ce ne sont pas à proprement parler les réponses que vous allez me donner, mais le processus mental qui va vous mener à ces réponses. Pour comprendre ça, je vais vous demander de « penser à voix haute ». Ne vous inquiétez pas, des consignes précises vous seront données au fur et à mesure. Vous n'avez qu'à vous laisser guider par mes questions.

Présentation de l'ESG



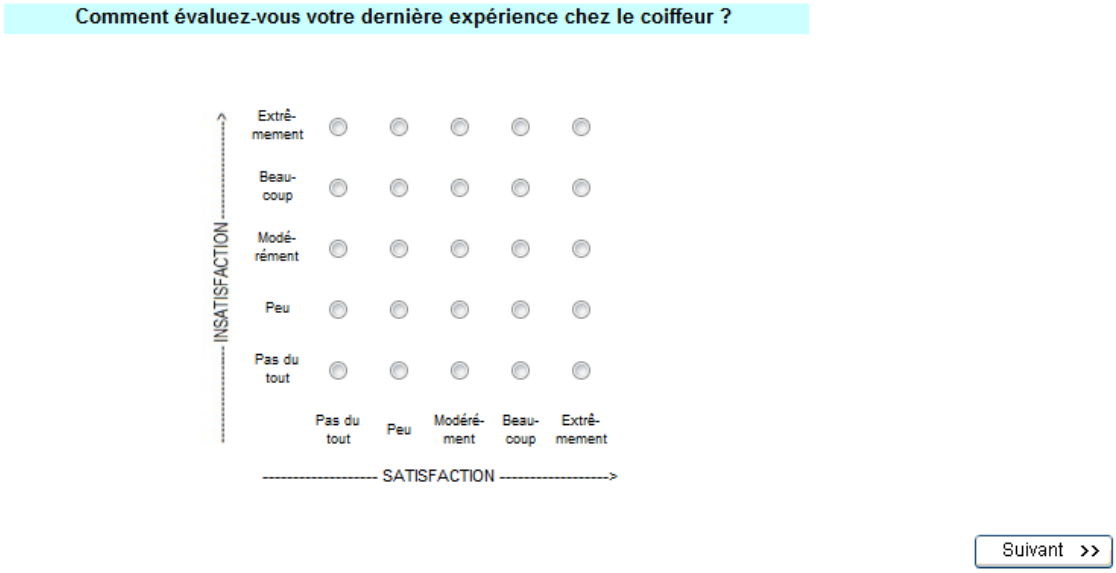
Pouvez-vous regarder attentivement cet outil de mesure et me dire comment vous le comprenez… ?

Pouvez-vous illustrer votre explication en me disant comment vous procèderiez si l'on vous demandait d'évaluer votre banque par exemple avec cet outil de mesure ?

Explication fournie en cas d'incompréhension de l'utilisation de l'ESG par la personne interrogée
Cette grille fonctionne de la manière suivante : Sur l'axe horizontal, on mesure le niveau de satisfaction (qui peut aller de « pas du tout » à « extrêmement ») ; Sur l'axe vertical, on mesure le niveau d'insatisfaction (qui peut aller de « pas du tout » à « extrêmement ») ; Ainsi, la pastille choisie (en cliquant dessus) correspond à la synthèse entre votre niveau de satisfaction et votre niveau d'insatisfaction.

Mise en situation d'évaluation

Maintenant pouvez-vous me dire ce qui se passe dans votre tête lorsque je vous mets face à cette grille pour évaluer votre dernière expérience chez le coiffeur.

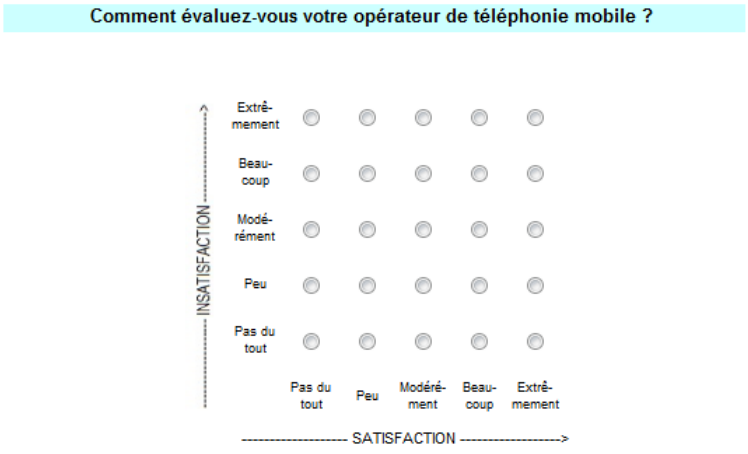


J'aimerais savoir précisément :

- Ce que vous regardez en 1er quand vous répondez.
- Toutes les étapes qui se produisent dans votre tête avant que vous ne choisissiez la pastille sur laquelle vous cliquez.

- Les raisons de ce niveau de satisfaction et de ce niveau d'insatisfaction.

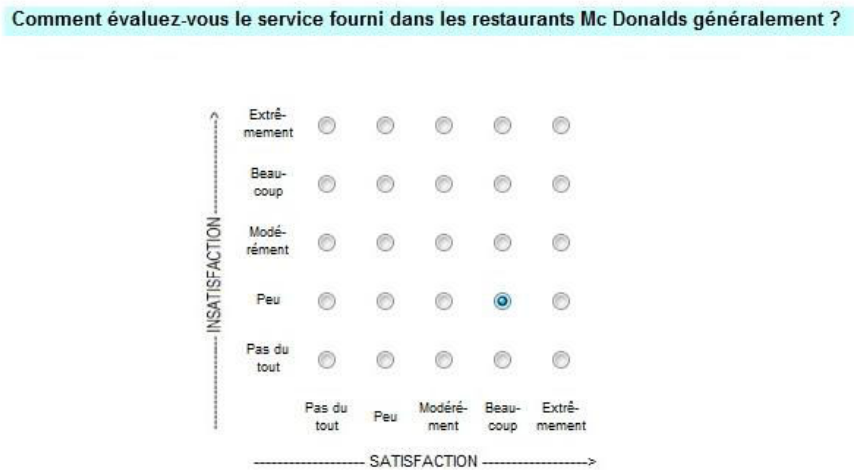
Même question pour l'opérateur de téléphonie mobile *(ou inversement, l'ordre d'apparition des pages du questionnaire étant aléatoire).*



Au final, est-ce que vous trouvez cet outil facile d'utilisation ? Intéressant ? Si oui, dans quelles situations ?

Interprétation des évaluations associées aux différentes cellules de l'ESG

Je vais maintenant vous montrer des exemples de réponses données par d'autres gens concernant différents services. J'aimerais savoir comment vous comprenez / interprétez ces réponses.

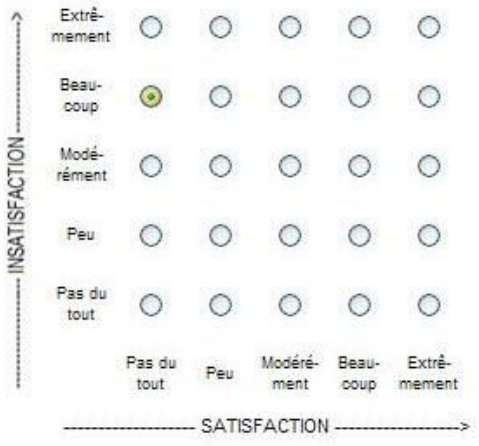


a. Comment décririez-vous la manière dont cette personne évalue les restaurants Mc Donald's ?
b. Pouvez-vous imaginer des exemples concrets de situations qui ont pu inciter la personne qui a répondu à cette question à répondre de cette manière ?

Les différentes zones de la grille sont ainsi explorées avec 5 évaluations différentes (une évaluation plutôt satisfaite, une évaluation plutôt insatisfaite, une évaluation indifférente et deux évaluations ambivalentes, l'une modérément, l'autre extrêmement). L'ordre d'apparition des pages est aléatoire.

Correspondance des évaluations de l'ESG sur un DS

La dernière partie de l'entretien vise à mettre en parallèle des évaluations effectuées avec cette grille avec des évaluations effectuées sur un autre outil de mesure que l'on utilise habituellement, une échelle classique en 9 points. A chaque fois, essayez de justifier le plus possible votre réponse.



Selon vous, à quoi cette évaluation correspond-t-elle sur cette échelle ?

INSATISFACTION ○ ○ ○ ○ ○ ○ ○ ○ ○ SATISFACTION

Les 16 des 25 possibilités de réponse sur la grille apparaissent aléatoirement.

MERCI !

Annexe 5 : Guide d'entretien de l'étude 4
(Deuxième partie : qu'en pensent les professionnels ?)

Introduction de l'entretien

Dans ma thèse j'ai travaillé sur la validation d'un point de vue académique d'un outil de mesure de l'ambivalence des consommateurs. J'arrive au terme de ma recherche et j'aimerais ouvrir mon travail au regard de professionnels pour estimer si cet outil pourrait concrètement avoir sa place dans la conduite d'études consommateurs.

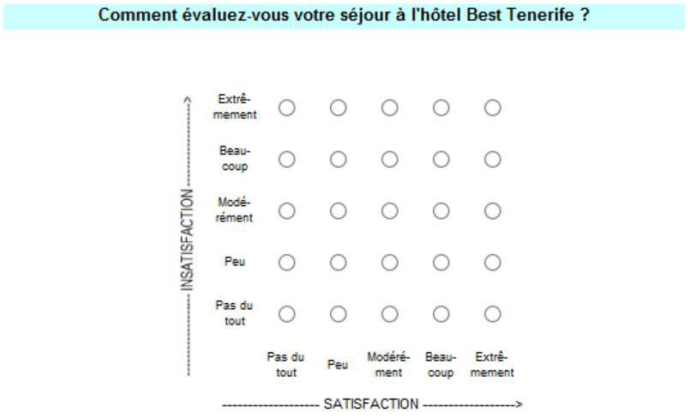
Mesure des réactions globales de consommateurs

Avant de vous montré l'outil de cet outil de mesure, j'aimerais comment vous mesurer généralement les réactions de consommateurs.

- Quel est le format des outils de mesure utilisés ?
- Est-ce que vous prenez en compte l'ambivalence d'une quelconque manière ?
- Est-ce qu'il vous paraitrait important de prendre en compte cet aspect ?

Réactions face à l'*Evaluative Space Grid*

Je vous envoie un lien vers un questionnaire en ligne utilisant l'outil de mesure que j'ai étudié dans ma thèse. Je vous laisse en prendre connaissance (exemple de question.



(Copie d'écran du questionnaire envoyé durant la conversation téléphonique)

Pouvez-vous le regarder attentivement et me dire comment vous le comprenez… ?

Quelles sont vos réactions vis-à-vis de cet outil ?

- Facilité d'utilisation ?
- Intéressant pour vos études ?
- Dans quel contexte précisément ?

Selon, vous cet outil offre-t-il une plus value par rapport à un différentiel sémantique classique ?

Annexe 6 : Questionnaire de l'étude 5

Page 1

Bonjour,

Vous allez participer une étude portant sur l'évaluation de votre médecin généraliste.

Lisez chaque question attentivement et prenez le temps de bien NUANCER vos réponses.

Les réponses sont strictement anonymes et confidentielles. Il n'y a pas de bonnes ou de mauvaises réponses, c'est toujours votre OPINION PERSONNELLE et SUBJECTIVE qui nous intéresse.

Merci d'avance pour votre participation !

Cliquez sur le bouton "suivant" pour commencer, svp. Suivant >

Page 2

QUELQUES QUESTIONS SUR VOUS POUR COMMENCER

Quel âge avez-vous ?

- ○ Entre 18 et 29 ans
- ○ Entre 30 et 40 ans
- ○ Entre 41 et 50 ans
- ○ Plus de 50 ans

Vous êtes :

- ○ Un homme
- ○ Une femme

Suivant >

Page 3

Des questions sur votre médecin généraliste* vont suivre.

Prenez quelques instants pour penser à votre relation avec celui-ci.

Cliquez sur le bouton "suivant" quand vous serez prêt à répondre à des questions sur ce sujet.

** Concentrez-vous sur le médecin généraliste que vous consultez le plus souvent*

Suivant >

Page 4

VOUS ET VOTRE MEDECIN GENERALISTE...

Depuis combien de temps consultez-vous ce médecin généraliste ?

- ◯ Depuis 0 à 3 mois
- ◯ Depuis 4 à 6 mois
- ◯ Depuis 7 mois à 1 an
- ◯ Depuis 1 à 2 ans
- ◯ Depuis 3 à 5 ans
- ◯ Depuis + de 5 ans

Votre relation avec votre médecin généraliste, pour vous c'est :

C'est sans importance	◯	◯	◯	◯	◯	◯	◯	C'est très important
Cela ne compte pas pour moi	◯	◯	◯	◯	◯	◯	◯	Cela compte beaucoup pour moi
Cela ne me concerne pas	◯	◯	◯	◯	◯	◯	◯	Cela me concerne beaucoup
C'est négligeable	◯	◯	◯	◯	◯	◯	◯	C'est considérable
C'est inutile	◯	◯	◯	◯	◯	◯	◯	C'est utile

Comment qualifieriez-vous votre expertise dans le domaine de la santé ?

J'ai très peu de connaissances dans le domaine de la santé	◯	◯	◯	◯	◯	◯	◯	J'ai beaucoup de connaissances dans le domaine de la santé
Je suis inexpérimenté(e) dans le domaine de la santé	◯	◯	◯	◯	◯	◯	◯	Je suis expérimenté(e) dans le domaine de la santé
Je ne suis pas informé(e) dans le domaine de la santé	◯	◯	◯	◯	◯	◯	◯	Je suis informé(e) dans le domaine de la santé
Je suis un novice dans le domaine de la santé	◯	◯	◯	◯	◯	◯	◯	Je suis un(e) expert(e) dans le domaine de la santé

Suivant >

Page 5

Dans la page qui va suivre, il va vous être demandé de nous fournir une évaluation globale de votre médecin généraliste.

Cette évaluation va s'effectuer sur une grille identique à celle que vous pouvez voir ci-dessous.

Votre opinions générale sera évaluée à l'aide de 3 grilles construites suivant le même schéma.

Sur cette grille :
- l'AXE VERTICAL vous permet d'indiquer dans quelle mesure vous êtes NEGATIF vis-à-vis de votre médecin généraliste (de "pas du tout" à "extrêmement")
- l'AXE HORIZONTAL vous permet d'indiquer dans quelle mesure vous êtes POSITIF vis-à-vis de votre médecin généraliste (de "pas du tout" à "extrêmement")

Vous pouvez ainsi cliquer sur la pastille qui décrit le mieux l'opinion générale (positive et négative) que vous vous faites de votre médecin généraliste.

EXEMPLE :
Si vous jugez que votre médecin généraliste vous fournit une prestation globalement mauvaise, mais qu' il y a tout de même quelques points positifs, vous pourrez cliquer sur la pastille suivante : (voir exemple ci-contre)



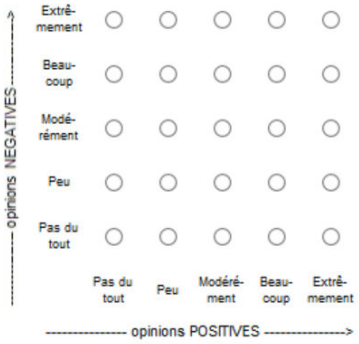
La réponse ci-dessus correspond donc à une évaluation :
- "beaucoup" négative
- et "peu" positive

Page 6

QUELLE EST VOTRE OPINION GENERALE SUR VOTRE MEDECIN GENERALISTE ?

Opinions négatives / opinions positives :

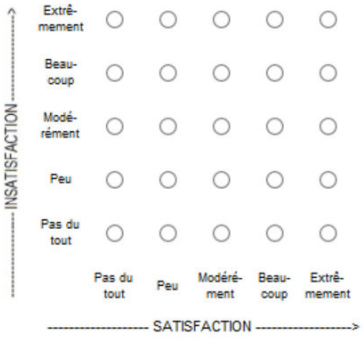
Pensez-vous que cette réponse reflète parfaitement votre opinion ?



------> *Pas du tout* ○ ○ ○ ○ ○ *Tout à fait*

Insatisfaction / satisfaction :

Pensez-vous que cette réponse reflète parfaitement votre opinion ?



------> *Pas du tout* ○ ○ ○ ○ ○ *Tout à fait*

Opinions défavorables / opinions favorables :

Pensez-vous que cette réponse reflète parfaitement votre opinion ?



------> *Pas du tout certain(e)* ○ ○ ○ ○ ○ *Absolument certain(e)*

< Précédent

Suivant >

Page 7

PLUS PRECISEMENT, VOTRE OPINION SUR VOTRE MEDECIN GENERALISTE...

Etes-vous d'accord avec les affirmations suivantes concernant votre médecin généraliste ?

Compétences médicales	Pas du tout d'accord						Tout à fait d'accord
Je pense que le cabinet de mon médecin généraliste dispose de tout le nécessaire pour dispenser des soins complets	○	○	○	○	○	○	○
Parfois, je me demande si les diagnostics de mon médecin généraliste sont corrects	○	○	○	○	○	○	○
Quand je consulte mon médecin généraliste, son examen est toujours complet	○	○	○	○	○	○	○
J'ai quelques doutes sur les compétences de mon médecin généraliste	○	○	○	○	○	○	○

Compétences relationnelles	Pas du tout d'accord						Tout à fait d'accord
Mon médecin généraliste se comporte de manière trop commerciale et impersonnelle avec moi	○	○	○	○	○	○	○
Mon médecin généraliste est amical et courtois avec moi	○	○	○	○	○	○	○

Capacités de communication	Pas du tout d'accord						Tout à fait d'accord
Mon médecin généraliste m'explique clairement les raisons des examens médicaux qu'il me fait passer	○	○	○	○	○	○	○
Parfois, mon médecin généraliste n'écoute pas ce que je lui dis	○	○	○	○	○	○	○

Temps passé avec le médecin	Pas du tout d'accord						Tout à fait d'accord
Mon médecin généraliste est parfois expéditif lors des consultations	○	○	○	○	○	○	○
Mon médecin généraliste passe en général suffisamment de temps avec moi	○	○	○	○	○	○	○

Accessibilité	Pas du tout d'accord						Tout à fait d'accord
Le cabinet de mon médecin généraliste est facilement accessible	○	○	○	○	○	○	○
Il y a beaucoup d'attente dans la salle d'attente de mon médecin généraliste	○	○	○	○	○	○	○
Je trouve qu'il est difficile d'obtenir un rendez-vous en temps voulu avec mon médecin généraliste	○	○	○	○	○	○	○
Il m'est possible de me faire soigner par mon médecin généraliste lorsque j'en ai besoin	○	○	○	○	○	○	○

Suivant >

Page 8

QUELQUES AUTRES QUESTIONS SUR VOUS POUR TERMINER

Quelle est votre année de naissance ?

Merci d'indiquer votre situation familiale :

- ◯ Seul(e) sans enfant
- ◯ Seul(e) avec enfant(s)
- ◯ En couple sans enfant
- ◯ En couple avec enfant(s)

Quel est votre dipôme le plus élevé ?

- ◯ Aucun diplôme
- ◯ Brevet des collèges
- ◯ CAP / BEP
- ◯ Baccalauréat général ou professionnel
- ◯ BTS / DUT
- ◯ Licence
- ◯ Maîtrise / Diplôme d'école de commerce ou d'ingénieur
- ◯ Master / DEA / DESS
- ◯ Doctorat

Valider

– Table des matières –

INTRODUCTION **11**

CHAPITRE 1. DEFINITION ET MESURE DE L'AMBIVALENCE **19**

SECTION 1. QU'EST-CE QUE L'AMBIVALENCE ? 19
1. L'ambivalence en psychologie *20*
2. L'ambivalence des consommateurs *22*
3. Les domaines d'étude de l'ambivalence *23*
4. Les propriétés associées à l'ambivalence *24*
5. Les conséquences comportementales de l'ambivalence *25*

SECTION 2. COMMENT MESURER L'AMBIVALENCE ? 26
1. Le différentiel sémantique, outil de mesure bipolaire *27*
2. Les limites du différentiel sémantique *29*
3. Les solutions proposées dans la littérature *31*
3.1.Les paires de continuums unipolaires *31*
3.2.Le calcul du score d'ambivalence *33*
4. L'approche bidimensionnelle *37*

SECTION 3. ETUDE PRELIMINAIRE SUR LA MESURE DES EVALUATIONS GLOBALES DE CONSOMMATEURS DANS LE E-COMMERCE 40
1. Panorama des outils utilisés dans le e-commerce pour mesurer et communiquer les évaluations globales de consommateurs *40*
1.1.Méthode *40*
1.2.Analyse et résultats *42*
1.3.Conclusion *48*
2. Sens attribué par les internautes aux points des dispositifs de mesure utilisés par les sites de e-commerce *48*
2.1.Méthode *49*
2.2.Analyse et résultats *50*
2.3.Conclusion *55*

CHAPITRE 2. L'*EVALUATIVE SPACE GRID*, UN OUTIL DE MESURE ALTERNATIF POUR SAISIR LA PART D'AMBIVALENCE DES CONSOMMATEURS 59

SECTION 1. L'EVALUATIVE SPACE GRID 60
1. Présentation de l'outil *60*
2. Validité de l'outil *62*
SECTION 2. APPLICATION DE L'*EVALUATIVE SPACE GRID* A LA MESURE DES EVALUATIONS DE CONSOMMATEURS 65
1. Délimitation de l'usage de l'outil *66*
1.1.Une mesure des réactions affectives et cognitives *66*
1.2.Une mesure mono-item *67*
1.3.Une mesure globale *67*
2. Résultats attendus *68*

CHAPITRE 3. VALIDITE CONVERGENTE DE L'*EVALUATIVE SPACE GRID* 79

SECTION 1. ANALYSE DE LA VALIDITE CONVERGENTE DE LA GRILLE AVEC LES PAIRES DE CONTINUUMS UNIPOLAIRES (ETUDE 1) 79
1. Méthode *80*
1.1.Design et stimuli *80*
1.2.Participants *82*
1.3.Procédure *82*
2. Analyses *84*
3. Résultats *85*
4. Conclusion *89*
5. Limites *89*
SECTION 2. ANALYSE DE LA VALIDITE CONVERGENTE DE LA GRILLE AVEC LES DIFFERENTIELS SEMANTIQUES, LES DISCOURS DES CONSOMMATEURS ET UNE MESURE DE L'AMBIVALENCE SUBJECTIVE (ETUDE 2) 91
1. Méthode *91*
1.1.Design et stimuli *91*
1.2.Participants *92*
1.3.Procédure *92*
2. Analyses *93*
2.1.Préparation des données *94*
2.2.Tests statistiques effectués *96*
3. Résultats *96*
3.1.Comparaison des évaluations issues de la grille avec le discours évaluatif des consommateurs *96*
3.2.Comparaison des évaluations issues de la grille de Larsen avec les notes obtenues sur les différentiels sémantiques *98*
3.3.Comparaison des évaluations issues de la grille de Larsen avec les scores moyens d'ambivalence subjective *99*
4. Conclusion *100*
5. Limites *101*

CHAPITRE 4. ETUDE DE LA PERTINENCE DE L'*EVALUATIVE SPACE GRID* COMME OUTIL DE MESURE DE LA SATISFACTION CLIENT (ETUDE 3) 103

SECTION 1. PRESENTATION DE L'ETUDE ET FORMULATION DES HYPOTHESES 104
SECTION 2. METHODE 105
1. Choix de l'objet d'étude *105*
2. Procédure *106*
3. Participants *106*
4. Format du questionnaire *107*
5. Echelles de mesure utilisées *107*

5.1.Satisfaction 108
5.2.Qualité perçue 108
5.3.Intentions comportementales 108
SECTION 3. ANALYSES ET RESULTATS 110
1. Procédure de codage préalable 110
2. Analyse des données 114
2.1.Analyses factorielles exploratoires 115
2.2.Analyses factorielles confirmatoires 120
2.3.Test du modèle structurel 123
SECTION 4. CONCLUSION 127
SECTION 5. LIMITES 128

CHAPITRE 5. PROCESSUS D'APPROPRIATION DE L'*EVALUATIVE SPACE GRID* PAR LES REPONDANTS : L'AVIS DES EXPERTS (ETUDE 4) 133

SECTION 1. QU'EN PENSENT LES CHERCHEURS ? 134
1. Présentation de l'étude et méthode 134
1.1.Nombre d'entretiens et choix de l'échantillon 134
1.2.Déroulement des entretiens 134
2. Analyses 136
3. Résultats 136
3.1.Prise en main de l'Evaluative Space Grid 136
3.2.Processus de formation de l'évaluation 138
3.3.Interprétation des différentes zones de l'Evaluative Space Grid 144
3.4.Intérêt perçu de l'Evaluative Space Grid 155
4. Conclusion 158
SECTION 2. QU'EN PENSENT LES PROFESSIONNELS ? 159
1.Présentation de l'étude et méthode 159
1.1.Nombre d'entretiens et choix de l'échantillon 159
1.2.Déroulement des entretiens 160
2. Analyses et résultats 161
2.1.La mesure des évaluations globales 161
2.2.Prise en compte de l'ambivalence 163
2.3.Réactions face à l'Evaluative Space Grid 164
2.4.Domaines d'application envisagés 166
SECTION 3. CONCLUSION 167

CHAPITRE 6. LA PLACE DE L'INCERTITUDE DANS LES EVALUATIONS EFFECTUEES AVEC L'*EVALUATIVE SPACE GRID* (ETUDE 5) 171

SECTION 1. CADRE THEORIQUE DE L'ETUDE 172
1. Un premier critère de distinction entre indifférence et ambivalence : le degré d'intensité des réactions 172
2. Un critère à prendre en compte pour affiner cette distinction : le degré de certitude 173
2.1.Qu'est-ce qu'une évaluation incertaine ? 173
2.2.Quels liens peut-on établir entre incertitude, indifférence et ambivalence ? 173
3. Combinaison des critères pour expliquer la répartition des évaluations sur la diagonale de la grille 175
SECTION 2. METHODOLOGIE 177
1. Choix de l'objet d'étude 177
2. Procédure et participants 177
3. Format du questionnaire 177

SECTION 3. ANALYSES ET RESULTATS 178
1. Codage préalable des données *178*
2. Analyse des données *178*
2.1.Dispersion des données *178*
2.2.Test de l'hypothèse H1 *179*
2.3.Test de l'hypothèse H2 *180*
SECTION 4. CONCLUSION 181
SECTION 5. LIMITES 183

CONCLUSION GENERALE 185

SECTION 1. CONTRIBUTIONS 186
1. Contributions méthodologiques *186*
2. Contributions théoriques *187*
3. Contributions managériales *188*
SECTION 2. LIMITES ET VOIES DE RECHERCHE 190
1. Vers d'autres opérationnalisations *190*
2. Vers l'étude de variables individuelles *193*

REFERENCES 195

LISTE DES TABLEAUX[15]

Tableau 1 - Scores d'ambivalence selon la formule de Kaplan (1972) 34

Tableau 2 - Scores d'ambivalence selon la formule de Katz et Hass (1988) 34

Tableau 3 - Scores d'ambivalence selon la formule de Jamieson (1988) (1988) 35

Tableau 4 - Scores d'ambivalence selon la formule de Thompson et al. (1995) 36

Tableau 5 - Corrélations entre les différentes mesures de l'ambivalence 37

Tableau 6 - Composition de l'échantillon de sites de e-commerce de l'étude préliminaire 41

Tableau 7 - Analyse descriptive des dispositifs de mesure proposés sur les principaux sites de e-commerce français 44

Tableau 8 - Présentation des cinq études empiriques de la thèse 76

Tableau 9 - Description de la variable « valence des commentaires ». 80

Tableau 10 - Résultats de l'ANOVA mixte : 2 (outil de mesure) x 3 (valence des commentaires) x 2 (continuum de mesure) ..... .86

Tableau 11 - Comparaisons de moyennes entre les conditions pour la note de négativité attribuée au corpus positif et pour la note de positivité attribuée au corpus négatif. ..... .88

Tableau 12 - Préparation des variables pour la comparaison entre les verbatims recodés et les notes issues de la grille de Larsen 94

Tableau 13 - Résultats des corrélations entre les verbatims recodés et les notes issues de la grille de Larsen 97

Tableau 14 - Résultats des ANOVA sur mesures répétées entre les verbatims recodés et les notes issues de la grille de Larsen 98

Tableau 15 - Résultats des corrélations entre les notes des différentiels sémantiques et de la grille de Larsen 98

Tableau 16 - Résultats des ANOVA sur mesures répétées entre les différentiels sémantiques et les notes issues de la grille de Larsen 99

Tableau 17 - Résultats des corrélations entre le score d'ambivalence subjective et le score d'ambivalence objective calculé à partir de la grille de Larsen 99

Tableau 18 - Résultats des différences de moyennes entre le score d'ambivalence subjective et le score d'ambivalence objective calculé à partir de la grille de Larsen 100

Tableau 19 - Moyenne et écart-type pour chaque item du questionnaire par condition 109

Tableau 20 - Impact des contraintes définies sur l'estimation des paramètres du modèle 113

[15] *Par défaut, les tableaux pour lesquels la source n'est pas indiquée sont source de l'auteur.*

Tableau 21 - Résultats suite à l'épuration de l'échelle de la qualité des banques.....118

Tableau 22 - Résultats des analyses exploratoires pour la satisfaction globale, les intentions de fidélité et les intentions de bouche-à-oreille.....119

Tableau 23 - Indicateurs de fiabilité et de validité convergente des construits.....120

Tableau 24 - Indices d'ajustement pour les modèles alternatifs testés entre la satisfaction globale et qualité perçue.....122

Tableau 25 - Indices d'ajustement pour les modèles alternatifs testés entre les intentions de fidélité et les intentions de bouche-à-oreille.....122

Tableau 26 - Indicateurs d'ajustement associés aux deux modèles de mesure.....124

Tableau 27 - Pourcentages de variance expliquée dans chaque ensemble de données.....124

Tableau 28 - Bêta standardisés et valeurs de t pour les effets directs et indirects de médiation suivant la méthode du *bootstrap* dans la condition DS.....126

Tableau 29 - Bêta standardisés et valeurs de t pour les effets directs et indirects de médiation suivant la méthode du *bootstrap* dans la condition ESG.....127

Tableau 30 - Profil des chercheurs interrogés.....134

Tableau 31 - Résumé des interprétations associées aux différents points de la diagonale.....155

Tableau 32 - Présentation des participants, de leur fonction et de leur entreprise.....160

Tableau 33 - Distinction entre indifférence et ambivalence suivant les critères d'intensité des réactions et de certitude dans la réponse donnée.....176

Tableau 34 - Test de différence de moyennes d'intention de bouche-à-oreille positif en fonction du type d'évaluation globale.....189

Tableau 35 - Comparaison des coefficients d'asymétrie pour les items de mesure de la satisfaction globale en fonction de l'outil de mesure utilisé (différentiel sémantique ou *Evaluative Space Grid*)..... 193

LISTE DES FIGURES[16]

Figure 1 – Réponses agrégées au niveau du point central des différentiels sémantiques 12

Figure 2 – Plan de la thèse 18

Figure 3 – La place de l'ambivalence dans le processus de consommation 22

Figure 4 - Interprétation des instructions du différentiel sémantique selon Kaplan (1972) 30

Figure 5 - La technique des composants ou des deux continuums unipolaires 32

Figure 6 - Le carré sémiotique des évaluations de consommateurs 39

Figure 7 - Question posée pour déterminer le sens donné aux différents points des échelles 49

Figure 8 - Analyse de la nature des dispositifs proposés par les répondants 51

Figure 9 - Représentation graphique de la distribution asymétrique négative de la satisfaction 52

Figure 10 - Analyse descriptive des étiquettes sémantiques proposées pour chaque point 54

Figure 11 - Traduction de l'*Evaluative Space Grid* 60

Figure 12 - Du différentiel sémantique à l'*Evaluative Space Grid*. 61

Figure 13 - Localisation des évaluations des différents types de stimuli sur la grille de Larsen 64

Figure 14 - Application de l'*Evaluative Space Grid* à la mesure de la satisfaction 65

Figure 15 - Localisation des quatre types de réactions évaluatives sur l'*Evaluative Space Grid* 66

Figure 16 - Opérationnalisation de l'*Evaluative Space Grid* pour la mesure des opinions (à gauche) et de la satisfaction globale (à droite) 73

Figure 17 - Design mixte séquentiel de la recherche 74

Figure 18 - Design de l'étude 1 81

Figure 19 - Item et outil de mesure utilisés pour évaluer les corpus de commentaires de consommateurs dans la *condition Evaluative Space Grid* 83

Figure 20 - Items et outils de mesure utilisés pour évaluer les corpus de commentaires de consommateurs dans la *condition deux continuums unipolaires* 83

Figure 21 - Format des données dans les deux conditions pour la variable « continuum de mesure » 84

Figure 22 - Codage de l'*Evaluative Space Grid* utilisé dans l'étude 1 85

Figure 23 - Scores moyens sur les continuums de la positivité et de la négativité pour les trois corpus de commentaires selon l'outil de mesure utilisé 87

Figure 24 - Application de la formule de Thompson et al. (1995) aux notes issues de la grille 96

[16] *Par défaut, les figures pour lesquelles la source n'est pas indiquée sont source de l'auteur*.

Figure 25 - Modèle théorique de la relation qualité perçue – satisfaction globale – intentions comportementales....104

Figure 26 - Transformation d'une grille de Larsen en différentiel sémantique en 9 points.... 111

Figure 27 - Codage de l'*Evaluative Space Grid* utilisé pour l'analyse 114

Figure 28 - Bêta standardisés et valeurs de t (entre parenthèses) pour le modèle structurel 125

Figure 29 - Arbre de codage des entretiens semi-directifs....137

Figure 30 - Matérialisation des différentes zones évaluatives de l'outil par les participants....154

Figure 31 - Dispositif de mesure utilisé par Vinivi (à gauche) et par TripAdvisor (à droite)....161

Figure 32 - Items utilisés par Vinivi pour l'évaluation d'un hôtel 164

Figure 33 - Proposition d'interprétation des différentes zones évaluatives de la grille de Larsen....168

Figure 34 - Ajustement de la relation entre les notes situées sur la diagonale de la grille et le niveau d'implication à l'égard du médecin suivant un modèle de régression linéaire et cubique....180

Figure 35 - Ajustement de la relation entre les notes situées sur la diagonale de la grille et la confiance qui leur est associée suivant un modèle de régression linéaire et quadratique.... 181

Figure 36 - Répartition des évaluations médianes le long de la diagonale de l'*Evaluative Space Grid*..... 182

Figure 37 - Répartition des zones évaluatives sur une *Evaluative Space Grid* en 9 points.... 191

LISTE DES ENCADRES

Encadré 1 - Détails sur l'article « *Are you Rather Satisfied or Dissatisfied? A Comparative Evaluation of Traditional and Alternative Scales for Measuring Overall Satisfaction* », projet développé avec Svein Ottar Olsen et Ana Alina Tudoran....128