

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

IDENTIFICATION, VALIDATION ET IMPORTANCE RELATIVE
DES DIMENSIONS DE QUALITÉ DES CHATBOTS TELS
QUE PERÇUS PAR LES UTILISATEURS

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN INFORMATIQUE DE GESTION

PAR
SEYF EDDINE GUESMIA

FÉVRIER 2019

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

TABLE DES MATIÈRES

LISTE DES FIGURES.....	vi
LISTE DES TABLEAUX.....	vii
RÉSUMÉ.....	ix
INTRODUCTION GÉNÉRALE.....	1
CHAPITRE I	
PROBLÉMATIQUE, OBJECTIF ET QUESTIONS DE RECHERCHE.....	5
1.1 Problématique	5
1.2 Objectif et questions de recherche	8
1.3 Étapes de réalisation.....	8
CHAPITRE II	
REVUE DE LA LITTÉRATURE.....	10
2.1 Introduction au chapitre	10
2.2 Intelligence artificielle et chatbots	11
2.3 Agents conversationnels.....	15
2.4 Systèmes QA (<i>Question Answering Systems</i>).....	16
2.4.1 Ambiguïté des questions	17
2.4.2 Architecture standard des systèmes de question-réponse	17
2.5 Concept de qualité.....	18
2.5.1 Définition du concept de qualité	18
2.5.2 Dimensions de la qualité	19
2.5.3 Qualité et Internet.....	20

2.5.4	Qualité et commerce électronique.....	21
2.5.5	Catégories de la qualité dans le commerce électronique	22
2.5.6	Qualité et chatbots.....	35
2.6	Conclusion au chapitre.....	38
CHAPITRE III		
	MODÈLE CONCEPTUEL ET APPROCHE MÉTHODOLOGIQUE.....	39
3.1	Introduction au chapitre	39
3.2	Modèle proposé.....	40
3.2.1	Catégorie « <i>Qualité de l'information</i> »	42
3.2.2	Catégorie « <i>Qualité du système</i> »	46
3.2.3	Catégorie « <i>Qualité de l'interaction</i> »	49
3.2.4	Autres Dimensions Reliées à la Qualité de Chatbots.....	50
3.3	Hypothèses	54
3.4	Validation et Fiabilité.....	54
3.4.1	Validation du contenu - Purification des items préliminaires du questionnaire	54
3.4.2	Validation du modèle	55
3.4.3	Fiabilité	56
3.5	Approche pour la collecte de données	57
3.5.1	Outil de collecte de données	57
3.5.2	Administration du questionnaire	58
3.5.3	Estimation de la grandeur de l'échantillon requis.....	58
3.6	Analyse des données	60
3.7	Conclusion au chapitre.....	60
CHAPITRE IV		
	ANALYSE ET PRÉSENTATION DES RÉSULTATS	61

4.1	Introduction au chapitre	61
4.2	Hypothèses	61
4.3	Profil démographique.....	62
4.4	Validation du modèle proposé	63
4.4.1	Raffinement du questionnaire à l'aide des juges.....	64
4.4.2	Raffinement du modèle en utilisant la méthode PLS-SEM	64
4.4.3	Spécifications Réflexives et Formatives	66
4.4.4	Description de notre modèle sous PLS-SEM.....	69
4.4.5	Grandeur de notre échantillon.....	72
4.5	Analyse du modèle externe	73
4.6	Analyse du modèle interne.....	79
4.7	Validité convergente du construit « <i>qualité du chatbot</i> »	84
4.8	Conclusion au chapitre	85
CHAPITRE V		
SOMMAIRE DES RESULTATS, IMPLICATIONS ET CONCLUSION		86
5.1	Sommaire des résultats.....	86
5.2	Implications théoriques	87
5.3	Implications pratiques	88
5.4	Implications managériales.....	89
5.5	Limites	89
5.6	Recherches futures	90
RÉFÉRENCES.....		93
ANNEXE A		
TABLEAU RÉCAPITULATIF DES FACTEURS DE QUALITÉ		
IDENTIFIÉS DANS LA REVUE DE LA LITTÉRATURE		104

ANNEXE B	
SIGNIFICATIONS DES FACTEURS DE	
QUALITÉ EXTRAITS DE LA LITTÉRATURE	108
ANNEXE C	
COURRIEL DE SOLLICITATION DES ENSEIGNANTS	116
ANNEXE D	
LETTRE DE CONSENTEMENT DES ENSEIGNANTS	118
ANNEXE E	
LETTRE D'INVITATION AUX PARTICIPANTS	120
ANNEXE F	
QUESTIONNAIRE.....	122
ANNEXE G	
CERTIFICAT D'APPROBATION ÉTHIQUE	125
ANNEXE H	
CERTIFICAT D'ACCOMPLISSEMENT (EPTC 2 : FER).....	127

LISTE DES FIGURES

Figures	Pages
2.1 Évolution des chatbots	13
3.1 Modèle proposé.....	41
4.1 Modélisations de Modèles Formatifs et Réflectifs	67
4.2 Modélisation combinant des aspects formatifs et réflexifs.....	69
4.3 Specification du modèle structurel initial sous SmartPLS.....	71
4.4 Modélisation de construit de 2e ordre à l'aide du « <i>Repeated Indicators Approach</i> » (Source : van Riel et coll., 2017).....	72
4.5 Modèle après avoir retirés deux items (Conv4 et Qual3Inv)	76
4.6 Nouveau modèle après avoir combiné les dimensions « <i>Complétude</i> » et « <i>Pertinence</i> »	78
4.7 Résultat du « <i>Bootstrapping</i> » montrant la signification statistique des « <i>paths</i> » du modèle.....	83

LISTE DES TABLEAUX

Tableaux	Pages
2.1 Exemples des questions d'ambiguïté	17
2.2 Dimensions de la qualité selon Garvin	20
2.3 Mesures de succès du commerce électronique du modèle de DeLone et McLean	23
2.4 Les dix différents aspects du fonctionnement de chaque assistant virtuel.....	25
2.5 Les cinq dimensions de SERVQUAL.....	26
2.6 Le cadre conceptuel de la qualité de l'interaction dans l'échange de services.....	29
2.7 Les dimensions de la qualité des sites Web	30
2.8 Échelle WebQual 4.0	32
2.9 Échelle eTailQ.....	33
2.10 Éléments de qualité (ou reliés à la qualité) pour les chatbots et les agents conversationnels (source Radziwill et Benton, 2017).....	36
3.1 Catégories, dimensions et items du modèle proposé	51
3.2 Les quatre chatbots utilisés lors de la collecte de données	58
3.3 Estimation de la taille de l'échantillon selon Marcoulides & Saunders.....	59
4.1 Profil démographique de l'échantillon.....	63
4.2 Dimensions après validation des experts	64
4.3 Les résultats de SmartPLS utilisés pour déterminer la fiabilité et la validité convergente	75

4.4	Valeurs des <i>Path Coefficients</i> générés par SmartPLS.....	80
4.5	Tableau récapitulatif – hypothèses et résultats.....	81

RÉSUMÉ

Les chatbots prouvent leur utilité dans l'automatisation des ventes et la délivrance de services de soutien à la clientèle. Leur capacité à interagir avec les clients dans un langage naturel, à répondre à leurs questions et à résoudre leurs problèmes permet aux entreprises d'économiser de l'argent en réduisant les coûts de main-d'œuvre. Il n'est pas surprenant que la demande des entreprises pour des chatbots bien conçus est en croissance. Cependant, peu d'études empiriques ont étudié les caractéristiques qui constituent un bon chatbot de service à la clientèle. L'objectif de notre étude est de combler cette lacune. Étant donné que nous cherchons à développer une mesure qui peut aussi servir comme outil de diagnostic pour des chatbots existants, nous avons exploré des études qui ont précédemment développé des mesures de qualité afin d'identifier des dimensions qui pourraient également être utiles pour refléter la qualité des chatbots. À partir de ça, nous avons identifié quinze dimensions de qualité. Ensuite, nous avons mené une enquête empirique. En utilisant un échantillon de $n = 87$ répondants pour valider ces dimensions et les classer en fonction de leur degré d'importance. Nous avons développé un questionnaire mis en ligne pour faire des analyses. Cette analyse permet d'identifier parmi les dimensions initialement proposées lesquelles sont reliées positivement et significativement à la qualité des chatbots. L'analyse a été faite en utilisant la méthode PLS-SEM (*Partial Least Square Structural Equation Modeling*) et le logiciel SmartPLS version 3. Les résultats montrent que, à partir des quinze dimensions identifiées, la qualité de l'information, la qualité du système, la qualité de l'interaction, l'apparence visuelle et le degré d'innovation du chatbot sont des dimensions de la qualité du chatbot (dans cet ordre d'influence).

Mots clés : chatbot, facteur de qualité, agent conversationnel, modèle de qualité.

INTRODUCTION GÉNÉRALE

Wikipédia définit un chatbot comme un programme informatique ou une intelligence artificielle qui peut mener une conversation par des méthodes auditives ou textuelles¹. De plus en plus, les chatbots démontrent leur capacité à améliorer l'expérience des clients dans le domaine du commerce électronique (Joshi et coll., 2017). C'est pour cela que les entreprises qui font des affaires en ligne portent un grand intérêt à expérimenter avec ces technologies. KLM Royal Dutch Airlines, par exemple, a conçu un chatbot en utilisant l'application Facebook Messenger. Appelé « BB » (pour BlueBot), le chatbot aide les passagers à réserver des billets et à les tenir au courant de l'état des vols et des changements de portes dans les aéroports. Durant les six premiers mois d'opération, ce chatbot a communiqué plus de deux millions de fois avec plus de 500 000 clients². Selon une étude, le marché des chatbots valait 864,9 millions USD en 2017 et atteindra 3 146,4 millions USD d'ici 2023. Cela représente un taux de croissance annuel moyen prévu de 24,1 % sur une période de 5 ans³.

Malgré cette popularité grandissante, les chatbots sont souvent confrontés à des problèmes liés à la qualité (Venkatesh et coll., 2017) et l'évaluation de celle-ci demeure encore difficile étant donnée l'absence de mesures de qualité de chatbots par des

¹ Wikipédia, <https://en.wikipedia.org/wiki/Chatbot>. Dernière visite : 8 Février 2019.

² Jay B., (2018), <https://www.convinceandconvert.com/digital-marketing/6-critical-chatbot-statistics-for-2018/>. Dernière visite : 8 août 2018.

³ Anonyme, (2018), https://www.mordorintelligence.com/industry-reports/chatbots-market?gclid=CjwKCAjwhqXbBRAREiwAucoo-6-7rz6ystOIVC8q4eU86OwpilPZT07DQGAgt-k5BiqcUrEBXaS6-BoCXwoQAvD_BwE. Dernière visite : 8 août 2018.

chercheurs (Kuligowska, 2015). En effet, un chatbot mal conçu peut facilement transformer une communication potentiellement positive avec un client en mauvaise expérience pour l'utilisateur, ce qui peut se refléter négativement sur une compagnie. L'expérience récente de Microsoft en est un exemple. La compagnie a lancé un chatbot nommé « Tay » pour les utilisateurs de Twitter. Tay était censé communiquer avec des personnes âgées entre 18 et 24 ans pour apprendre à mieux interagir avec ceux-ci en imitant leur façon de parler. Malheureusement, cet exercice a été interrompu par la compagnie lorsque Tay a commencé à envoyer des phrases vulgaires et offensantes sur Twitter, ce qui a nui à la renommée de Microsoft⁴.

Le concept de qualité a été défini et exploré pour la délivrance de services dans des contextes de commerce traditionnel et de vente en ligne depuis déjà plusieurs années (par ex., Parasuraman et coll., 1985, 1988, 1991 ; Zeithaml et coll., 1985 ; Loiacono et coll., 2002 ; Elling et coll., 2012). Néanmoins, peu de recherches ont exploré le concept de qualité pour les chatbots, et à notre connaissance, aucune recherche à ce jour, n'a proposé de développer une mesure de qualité multidimensionnelle qui pourrait être applicable spécifiquement aux chatbots comme plusieurs qui ont été développées pour évaluer les sites Web commerciaux (par ex. WebQual).

Pour les compagnies qui ont déjà une présence en ligne, il y a plusieurs avantages liés aux mesures de qualité pour l'évaluation de sites Web, lesquelles sont basées sur les attentes et perceptions des consommateurs et validées empiriquement. Une mesure de qualité permet non seulement d'évaluer l'attrait et le bon fonctionnement de sites Web existants, mais elle aide aussi à guider les développeurs de sites Web lors de la conception de nouveaux sites (Loiacono et coll., 2002). Bien que plusieurs critères

⁴ Carlo Cadet, (2018), <http://www.brandquarterly.com/battle-bots-quality-assurance-crucial-chatbot-success>. Dernière visite : 8 août 2018.

servant à évaluer la qualité des sites Web sont aussi vraisemblablement applicables aux chatbots (par ex. la facilité d'usage), le fonctionnement des chatbots et leur raison d'être différent de ceux des sites Web. Le chatbot est une interface interactive capable de recevoir des requêtes non structurées (questions posées par des clients), de les comprendre, de chercher les informations pertinentes dans les bases de données des compagnies, et ensuite de générer des réponses en langage naturel (Kuligowska, 2015). Un chatbot de haute qualité devrait afficher un comportement similaire à celui d'un être humain. Certains seraient même capables d'interpréter les émotions humaines, de traiter avec la colère des clients, de diagnostiquer des situations critiques que vivent certains clients avec les entreprises, et ensuite de proposer des solutions (Radziwill et Benton, 2017). Par conséquent, bien que certains critères de qualité des chatbots commerciaux (appelées aussi « *dimensions de qualité* ») sont probablement semblables à ceux déjà identifiés par les chercheurs qui se sont intéressés, comme déjà dit, au concept de qualité de service et de sites Web, l'aspect et le fonctionnement unique des chatbots warrants that the concept of quality différent et requièrent qu'un modèle de qualité spécifique évaluant la qualité des chatbots doit être développé (Parasuraman et coll., 1988).

Notre étude consiste à développer un modèle multidimensionnel de qualité qui peut être utilisé pour évaluer les chatbots utilisés dans des contextes commerciaux (voir Chapitre I). Pour cela, nous nous basons sur une revue de la littérature sur le concept de la qualité pour identifier les dimensions de qualité qui peuvent influencer la qualité du chatbot (voir Chapitre II). Nos résultats (voir Chapitres IV et V) montrent que la qualité de l'information, la qualité du système, la qualité d'interaction, l'apparence visuelle et le degré auquel le chatbot est perçu comme étant innovant sont des dimensions de la qualité perçue du chatbot (dans cet ordre d'importance).

La nature multidimensionnelle de notre modèle de qualité des chatbots (voir Chapitre III) permet non seulement de guider des concepteurs lors de la conception de nouveaux

chatbots mais aussi comme outil de diagnostic pour évaluer des chatbots existant afin d'améliorer la perception de leur qualité aux yeux des usagers.

Dans le chapitre I de notre étude, nous détaillons le sujet que nous traitons ainsi que le problématique et les objectifs soulignés. Dans le chapitre II, nous présentons une revue de la littérature sur l'intelligence artificielle, les chatbots et le concept de qualité. Nous y présentons des études qui ont cherché à identifier les dimensions de la qualité dans plusieurs domaines de recherche. Nous présentons ensuite notre modèle de recherche et nos hypothèses dans le chapitre III. Dans le chapitre IV, nous décrivons la procédure utilisée pour tester et valider notre modèle de recherche et nos hypothèses. Le chapitre V conclut notre rapport avec une discussion des résultats obtenus et les implications de cette recherche.

CHAPITRE I

PROBLÉMATIQUE, OBJECTIF ET QUESTIONS DE RECHERCHE

1.1 Problématique

L'influence des chatbots dans le quotidien des gens devient de plus en plus évidente. Par exemple, en Chine, des chatbots comme WeChat sont maintenant utilisés pour prendre des rendez-vous médicaux, transférer de l'argent à des amis, payer des factures, réserver des vols d'avion, etc. Plusieurs entreprises investissent dans les chatbots afin de faciliter leurs activités commerciales et répondre aux demandes de leurs clients⁵.

Le service à la clientèle est un secteur idéal pour l'application des chatbots. Appelés aussi « *agents conversationnels* », les chatbots acceptent les questions des clients et fournissent les informations nécessaires en utilisant un langage naturel. Pour les utilisateurs, les chatbots présentent plusieurs avantages. Parmi les avantages les plus souvent mentionnés, l'on retrouve la vitesse et la disponibilité. Précisément, 64 % des personnes interrogées ont déclaré qu'avoir accès aux services d'une compagnie 24

⁵ Anonyme, (2018), https://www.mordorintelligence.com/industry-reports/chatbots-market?gclid=CjwKCAjwhqXbBRAREiwAucoo-6-7rz6ystOIVC8q4eU86OwpilPZT07DQGAgt-k5BiqcUrEBXaS6-BoCXwoQAvD_BwE. Dernière visite : 8 août 2018.

heures sur 24 était l'avantage des chatbots. Le deuxième avantage le plus mentionné (par 55 % des participants) était la capacité d'obtenir des réponses instantanées⁶.

Un chatbot est un logiciel qui utilise l'intelligence artificielle et qui permet une conversation en utilisant des canaux de communications auditifs ou textuels. De tels programmes sont souvent conçus pour simuler de manière convaincante comment un humain se comporte comme partenaire de conversation⁷. Étant donné que les chatbots peuvent interagir comme des humains, les entreprises les choisissent pour communiquer avec leurs clients. Les chatbots affichent des comportements humains tels que sourire, plaisanter, se montrer poli, etc. En outre, grâce aux possibilités offertes aujourd'hui par la technologie, les chercheurs sont persuadés que les agents conversationnels deviendront de plus en plus intelligents en raison de leur capacité grandissante à performer des fonctionnalités plus compliquées, comme le diagnostic de problèmes complexes (Kuligowska, 2015; Radziwill et Benton, 2017). Cette fonctionnalité rend les chatbots capables d'aider à mieux comprendre les problèmes des clients tout en diagnostiquant leurs situations et en proposant des solutions. C'est une bonne raison pour laquelle plusieurs entreprises croient fermement que ce sera le moyen idéal de commercialiser leurs produits et services, et de résoudre des défaillances de service (Xu et coll., 2017).

Notre revue de la littérature sur les chatbots montre que la plupart des études menées à ce jour mettent une priorité sur la présentation d'améliorations au niveau technologique ; cela comprend l'usage de langages naturels, l'extraction de mots clés à partir des requêtes formulées par les clients, la recherche d'informations dans les

⁶ Jay B., (2018), <https://www.convinceandconvert.com/digital-marketing/6-critical-chatbot-statistics-for-2018/>. Dernière visite : 8 août 2018.

⁷ Wikipédia, <https://en.wikipedia.org/wiki/Chatbot>. Dernière visite: 9 août 2018.

bases de données des entreprises, etc. Par contre, peu d'études ont été menées pour définir le concept de qualité dans le contexte des chatbots et essayer d'identifier les critères ou facettes utilisés par les utilisateurs afin d'évaluer la qualité des chatbots.

Ces critères sont souvent appelés « *dimensions* » dans la littérature sur la qualité perçue et dans d'autres recherches dans les sciences sociales. Dans cette littérature, des concepts tels que la qualité sont typiquement décrits comme étant des « *construits* ». Dans la littérature sur les méthodologies quantitatives, les construits représentent des abstractions mentales utilisées pour exprimer des idées et des concepts qui nous intéressent. En d'autres mots, un construit est une idée abstraite que l'on souhaite mesurer à l'aide d'une échelle dans un questionnaire.

Certains construits sont simples ; on peut les mesurer en utilisant seulement une ou quelques questions. D'autres construits sont plus complexes, par exemple la qualité perçue. Les construits complexes comportent des dimensions (facettes) multiples qui sont liées de manière commune et composent le construit qu'elles reflètent. Dans un modèle représentant un concept multidimensionnel, les relations entre le construit et ses dimensions ne sont pas des forces causales, mais elles représentent plutôt des associations entre le concept général qui nous intéresse et les dimensions qui *reflètent* ou constituent sa construction (Law et coll., 1998).

À notre connaissance, aucune recherche n'a tenté de développer une mesure de qualité qui se base sur plusieurs dimensions de la qualité perçue des chatbots commerciaux. Par conséquent, peu d'informations existent pour aider les développeurs à identifier les dimensions de qualité qu'ils devraient prendre en compte pour développer de nouveaux chatbots de haute qualité. De plus, pour les compagnies qui ont déjà implémenté des chatbots, le manque de mesure multidimensionnelle de qualité perçue veut dire qu'il n'y a pas d'instruments de diagnostic qui pourraient être utilisés afin d'améliorer leur qualité. Pour cela, nous croyons qu'une mesure de qualité spécialement conçue pour

évaluer la qualité des chatbots représente une contribution importante dans le domaine. Ci-dessous, nous présentons notre question de recherche et les objectifs de notre étude.

1.2 Objectif et questions de recherche

L'évaluation de la qualité des chatbots se base sur des facettes qui reflètent différents aspects : par exemple leur raison d'être, l'information que les chatbots produisent, leur fonctionnement efficace, la fiabilité et le bon fonctionnement des technologies sur lesquelles les chatbots s'appuient. Ces facettes ou dimensions peuvent englober des aspects informationnels, des aspects technologiques, mais aussi des aspects interactionnels entre le chatbot et l'utilisateur.

Notre objectif général est d'identifier les dimensions de la qualité des chatbots en nous basant sur les perceptions des utilisateurs. Plus précisément, les questions de recherche suivantes guideront notre étude :

- **Quelles sont les dimensions qui reflètent les perceptions de qualité d'un chatbot ?**
- **Quelle est l'importance relative de chacune de ces dimensions comparativement les unes aux autres ?**

1.3 Étapes de réalisation

- Étape 1 – Identifier, à travers une revue de la littérature, les dimensions potentielles de la qualité perçue par les utilisateurs de chatbots dans des contextes commerciaux.
- Étape 2 – Proposer un modèle de qualité qui conçoit la qualité perçue des chatbots comme un construit multidimensionnel et le tester / valider empiriquement à l'aide du logiciel SmartPLS (version n° 3).

- Étape 3 – Nettoyer le modèle préliminaire proposé à l'étape n° 2 en le réduisant à ses dimensions les plus importantes. L'importance de chaque dimension de qualité représentera le degré d'influence de cette dimension sur la perception des utilisateurs au niveau de la qualité des chatbots.

CHAPITRE II

REVUE DE LA LITTÉRATURE

2.1 Introduction au chapitre

Le but de notre projet est d'identifier et de valider les dimensions de qualité dans le contexte des chatbots utilisés pour le service à la clientèle. À notre connaissance, aucune étude publiée dans une revue scientifique n'a auparavant essayé de développer et valider empiriquement une mesure multidimensionnelle de qualité conçus spécialement pour les Chatbots.

Afin de guider notre revue de littérature, nous avons utilisé une approche méthodique pour collecter les articles. Cette approche se base sur deux critères majeurs : (1) l'année de la publication et (2) la pertinence de la publication relativement à notre étude. Nous avons reculé cinq ans en arrière pour collecter des articles contenant des contributions qui ont aidé à développer le domaine d'étude. En d'autres mots, nous avons cherché identifier les articles scientifiques et chapitres de livres qui ont paru entre 2013 et 2018. De plus anciennes publications contenant des œuvres originales considérées comme la base du sujet étudié (« les classiques ») ont eux aussi été sélectionnées pour notre revue de littérature.

Pour la pertinence de l'étude, premièrement, les publications ont été sélectionnées si elles contenaient au moins un des termes de recherche identifiés dans le titre et / ou le résumé afin d'assurer la pertinence de la collecte de données. Deuxièmement, seules

les publications dont la qualité constituait un aspect majeur ou influent de la recherche ont été sélectionnées.

Les sections 2.2, 2.3 et 2.4 de ce chapitre présentent une revue de la littérature sur les agents conversationnels (chatbots). Nous chercherons principalement à définir ce que sont des chatbots, à expliquer brièvement leur évolution et à expliquer comment ils sont implémentés dans le contexte commercial. Finalement, nous identifions les types d'agents conversationnels couramment utilisés dans le secteur commercial. Les sections 2.5, 2.5.1, 2.5.2 et 2.5.3 présentent une revue de la littérature sur la qualité et Internet en général (section 2.5.3). Pour les cas particuliers du commerce électronique, nous les présentons dans les sections 2.5.4 et 2.5.5. Enfin, la section 2.5.6 discute des études antérieures qui traitent des chatbots en termes de qualité.

2.2 Intelligence artificielle et chatbots

Bayerque (2016) a défini l'intelligence artificielle comme étant l'intelligence présentée par les machines, les présentant comme des agents rationnels capables de percevoir leur environnement, d'interagir et de prendre des décisions. Selon ce même auteur, l'agent conversationnel est un ordinateur capable de simuler, de façon réaliste, l'interaction humaine. Entre les années 1964 et 1966, l'informaticien Joseph Weizenbaum a inventé le premier chatbot appelé Eliza. Ce projet était basé sur les résultats de tests faits en 1950 par l'informaticien Alan Turing. Les expériences de Turing consistaient à tester la capacité d'un ordinateur à montrer un comportement intelligent équivalent ou indiscernable de celui d'un humain. Turing a considéré le concept d'ordinateur interactif comme étant une machine capable d'imiter l'intelligence humaine (voir aussi Kolb et Millican, 2006).

La fin des années 1970 et le début des années 1980 ont vu la commercialisation de microordinateurs⁸ à des prix de plus en plus abordables ainsi que d'importants développements dans le domaine de l'interaction Humain-Machine. Celle-ci a été largement étendue lorsque les scientifiques ont introduit l'interface utilisateur graphique IUG (ou GUI pour *Graphical User Interface*) afin de simplifier l'utilisation de l'ordinateur et éviter la nécessité de l'apprentissage du code binaire par l'utilisateur normal pour travailler avec des ordinateurs. Ils ont commencé par l'écran (qui affiche des icônes et des boutons) et ils ont fini par utiliser un langage naturel pour communiquer avec la machine (Wilbert, 2002). En somme, il devenait possible pour tout le monde d'avoir accès aux chatbots à l'aide d'un ordinateur ou un appareil portable (le cas des smart phones et des tablettes aujourd'hui). Au début, les chatbots n'avaient d'autre but que d'interagir avec des utilisateurs à l'aide d'interactions simples. Mais avec des améliorations continues au niveau de l'intelligence artificielle, du traitement automatique du langage, et des technologies numériques (mobiles, par ex.), en plus de l'accessibilité à Internet, les chatbots peuvent supporter des applications variées dans le secteur du service à la clientèle, par exemple informer les clients sur les produits et services offerts, ainsi qu'à interagir avec eux en utilisant un langage naturel (Bayerque, 2016). La figure 3.1 montre les secteurs majeurs qui ont fait grande partie de l'évolution des chatbots dès le début de leur existence (voir Lalwani et coll., 2018).

⁸ Wikipédia, https://fr.wikipedia.org/wiki/Histoire_des_ordinateurs_personnels. Dernière visite : 29 août 2018.

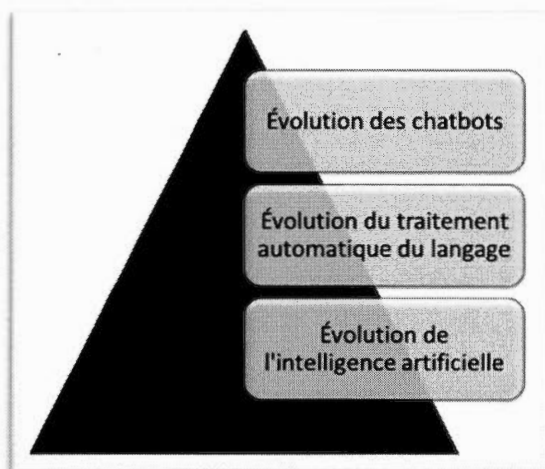


Figure 2.1 – Évolution des chatbots

En effet, en raison d'une concurrence de plus en plus mondiale, un nombre grandissant d'entreprises choisissent d'investir dans la technologie des chatbots afin d'améliorer leur service à la clientèle tout en réduisant leurs coûts. Selon une étude menée par Gartner, d'ici 2020, les chatbots traiteront pas moins de 85 % de toutes les interactions avec la clientèle⁹. Cela n'est pas surprenant étant donné que les chatbots offrent plusieurs avantages¹⁰, par exemple :

- Capacité à répondre aux questions des clients 24 heures sur 24, 7 jours par semaine.

⁹ CRM Strategies and Technologies to Understand, Grow and Manage Customer Experiences. (2011), https://www.gartner.com/imagesrv/summits/docs/na/customer-360/C360_2011_brochure_FINAL.pdf. Dernière visite : 30 août 2018.

¹⁰ Top 5 Benefits Of Using Chatbots For Your Business. (2017), <https://chatbotsmagazine.com/top-5-benefits-with-using-chatbots-for-your-business-159a0cee7d8a>. Dernière visite : 30 août 2018.

- Traitement d'un plus grand nombre des demandes de la clientèle en moins de temps¹¹.
- Réduction du personnel, et donc des coûts.
- Distribution des ressources humaines à l'intérieur de l'organisation de manière plus intelligente en automatisant des tâches routinières.

Si l'on prend, par exemple, le secteur bancaire. Des tâches routinières (telles que la demande de solde, les détails des comptes bancaires, des demandes de prêt, etc.) peuvent aisément et efficacement être traitées aujourd'hui par des chatbots sans avoir besoin que le client se déplace physiquement ou ait à attendre au téléphone afin de parler à un représentant humain. Cela permet aux banques de mieux gérer leur service à la clientèle en réassignant leur personnel humain à résoudre des problèmes plus complexes qui nécessitent une interaction humaine et personnalisée. En somme, les chatbots et les nouvelles technologies numériques telles que le mobile promettent d'apporter des économies annuelles importantes. Selon un rapport publié par Juniper Networks, cette figure pourrait atteindre plus de huit milliards de dollars américains d'ici 2022 uniquement dans le secteur bancaire¹².

¹¹ Les êtres humains sont limités par le nombre d'activités qu'ils peuvent exécuter en un temps donné, sinon le risque d'erreurs augmente. Les chatbots adressent cette limite humaine directement promettant à améliorer le service à la clientèle. De plus, si l'on utilise des chatbots pour répondre aux questions des clients, il n'y a plus besoin de mettre les clients en attentes. Ils peuvent interagir instantanément avec un chatbot lorsqu'ils communiquent avec la compagnie.

¹² Smith, S. (2016). Mobile Banking Users to Reach 2 billion by 2020, Representing more than 1 in 3 of Global Adult Population. <https://www.juniperresearch.com/press/press-releases/mobile-banking-users-to-reach-2-billion-by-2020>. Dernière visite : 30 août 2018.

2.3 Agents conversationnels

La façon traditionnelle d'interagir avec un site Web est de sélectionner manuellement des chemins logiques préfigurés pour celui-ci sous la forme de menus ou de sous-menu correspondant à des utilisations déterminées d'avance. Lorsque la tâche à réaliser est complexe, ou que l'utilisateur ne connaît pas bien la terminologie du domaine, cette méthode traditionnelle de navigation n'offre pas toujours une interface efficace. C'est dans ces cas, où l'utilisateur peut bénéficier d'agents conversationnels. Ces agents se distinguent par trois éléments : (1) ils s'alimentent du langage naturel ou pseudo naturel, (2) ils répondent en langage naturel, et (3) ils ne contiennent pas de chemins de navigation prédéterminés (Kuligowska, 2015). Les agents conversationnels peuvent comprendre plusieurs types de technologies, à savoir les systèmes de question-réponse, les chatbots, les robots sociotechniques, etc (Rahman, et coll., 2017). Dans notre étude, nous les étudions du point de vue de leur usage et du point de vue de leur stratégie d'implantation (Radziwill et Benton, 2017).

Selon Daeninck (2002), Les agents conversationnels sont des systèmes logiciels qui imitent les interactions humaines pour transmettre des informations de façon naturelle. Ils ont été conçus pour être en mesure d'utiliser différentes techniques similaires à celles que les humains utilisent, par exemple le typage de texte naturel, les mouvements du corps et les traits du visage (Daeninck, 2002).

Far (2003) affirme que les agents conversationnels doivent avoir quatre composantes communes : (1) une forme de saisie (il peut s'agir de saisie de texte, de reconnaissance vocale ou d'une combinaison des deux) ; (2) un analyseur de langage naturel (pour interpréter ce que dit l'utilisateur) ; (3) une forme de prise de décision ; et (4) une forme de sortie (pour convertir la décision prise en langage naturel). Ces composants sont le minimum nécessaire pour avoir un agent conversationnel. D'autres peuvent être ajoutés pour rendre l'agent plus semblable aux humains (Far, 2003).

Il existe de nombreux champs où l'on trouve l'utilité d'utiliser les agents conversationnels. Selon Daeninck (2002), les robots pourraient être utilisés comme des représentants commerciaux sur le Web, ou des guides Web pour trouver des sites ou des informations demandées par les utilisateurs. Ils pourraient être utilisés pour répondre aux questions les plus fréquemment posées. On peut même les utiliser comme des remplacements de services téléphoniques tactiles. Cette liste est loin d'être complète, car ces agents peuvent être utilisés pour toute tâche nécessitant une interaction avec l'utilisateur.

2.4 Systèmes QA (*Question Answering Systems*)

L'étape clé vers la réalisation de l'interaction intelligente entre l'humain et la machine est de permettre aux systèmes de questions-réponses de comprendre les besoins d'information de l'utilisateur (Li, 2011).

Un système de question-réponse est un type de système de recherche d'informations qui nécessite des techniques de traitement du langage naturel plus complexes. C'est un système capable d'obtenir des réponses à des questions posées en langage naturel à partir de documents disponibles sur Internet (Hirschman et Gaizauskas, 2001).

Les systèmes de question-réponse peuvent ne pas toujours répondre aux besoins d'information des utilisateurs. L'unité responsable du traitement peut ne pas catégoriser correctement la question, et parfois les informations nécessaires pour extraire et générer la réponse ne sont pas facilement récupérées (Burger et coll., 2001).

2.4.1 Ambiguïté des questions

Le besoin d'information pour un utilisateur peut être exprimé de diverses manières en raison de l'inhérence de l'ambiguïté dans le langage naturel. C'est la raison pour laquelle l'ambiguïté des questions dans les systèmes de question-réponse a été reconnue comme étant un problème critique pour l'interprétation automatique des expressions du langage naturel et le traitement des questions (Burger et coll., 2001 ; Unger et Cimiano, 2011). Kamsties (2001) a classé le problème de l'ambiguïté dans un langage naturel en quatre catégories : l'ambiguïté lexicale, l'ambiguïté structurale, l'ambiguïté sémantique et l'ambiguïté pragmatique. Le tableau 2.1 illustre plusieurs formes de ces ambiguïtés associées à quatre niveaux de sophistication de l'interrogateur (Kamsties, 2001).

Tableau 2.1 – Exemples des questions d'ambiguïté (Burger et coll., 2001)

Niveau 1	Q : Où Pfizer Pharmaceutical fait-elle des affaires ?	Ambiguïté – Granularité de localisation (pays, état, ville).
Niveau 2	Q : Quels médicaments récents Pfizer a-t-elle mis sur le marché ?	Ambiguïté – Médicaments développés par Pfizer ou commercialisés par Pfizer.
Niveau 3	Q : Quand le Viagra a-t-il été introduit sur le marché ? Q : Le Viagra a-t-il été un médicament efficace ?	Ambiguïté – Localisation du marché : mise en marché des médicaments aux États-Unis, en Europe ou au Mexique.
Niveau 4	Q : Pourquoi les entreprises pharmaceutiques utilisent-elles des applications cliniques basées sur Oracle ?	Ambiguïté – Impact sur la satisfaction du client, impact sur la productivité.

2.4.2 Architecture standard des systèmes de question-réponse

Fonctionnellement, la plupart des systèmes actuels de réponses aux questions peuvent être décomposés en quatre composantes : (1) l'analyse des questions, (2) la recherche des documents, (3) la recherche du passage, et (4) l'extraction des réponses (Tellex, 2003).

L'analyse des questions signifie convertir la question entrée en langage naturel, en requête structurée complexe pour être utilisée par d'autres modules (Tellex, 2003). La recherche des documents signifie de faire la recherche des documents du corpus susceptibles de contenir des réponses à la question de l'utilisateur (Hu, 2006). La recherche du passage signifie de chercher des blocs de plusieurs phrases contenant une réponse. Le système prend un document comme une question et essaie de trouver des passages du document contenant une réponse (Tellex, 2003). Finalement, l'extraction de la réponse signifie d'identifier les phrases de réponses exactes dans les passages recherchés. Le processus tente de récupérer une expression exacte pour la retourner comme une réponse en utilisant des algorithmes d'extraction de réponses (Yun, 2007).

2.5 Concept de qualité

2.5.1 Définition du concept de qualité

Durant de nombreuses décennies, le concept de qualité a été perçu différemment par plusieurs chercheurs. La raison de ce fait est que le concept a été utilisé dans plusieurs domaines. Parasuraman, Zeithaml et Berry (1988) (ceux qui ont développé la mesure de la qualité de service ou, tout simplement « *SERVQUAL* ») définissent la qualité comme étant l'écart entre les attentes et les perceptions de la performance actuelle. Ils décrivent la qualité du service perçue par les consommateurs comme étant le résultat d'une comparaison entre ce que le consommateur pense qu'une entreprise de services devrait offrir (ses attentes) et sa perception de la performance actuelle lors de la délivrance de services. En somme, lesdits auteurs considèrent la qualité de service perçue comme étant le degré et la direction de l'écart entre les perceptions et les attentes des consommateurs. Par exemple, si les attentes d'un client sont positives et très hautes par rapport à un service attendu, mais que ce client perçoit que ce qu'il a réellement

reçu est en-dessous de ses attentes, la qualité de service perçue sera perçue comme étant basse.

Parasuraman, Zeithaml et Berry (1988) ont utilisé le terme « *expectation* » (en français : attentes) dans leur littérature sur la qualité des services de façon différente de la façon dont ils l'ont utilisé dans leur littérature sur la satisfaction du consommateur. En particulier, dans la littérature de satisfaction du client, les attentes sont considérées comme étant des prédictions faites par les consommateurs sur ce qui se produira probablement lors d'une transaction ou d'un échange imminent. En revanche, dans la littérature sur la qualité des services, les attentes sont considérées comme étant les désirs des consommateurs, c'est-à-dire ce que ceux-ci pensent qu'un fournisseur de services devrait offrir, en d'autres termes, ce que le consommateur aimerait recevoir.

2.5.2 Dimensions de la qualité

Un construit multidimensionnel est un concept théorique qui est mesuré par plusieurs autres concepts connexes. En d'autres termes, c'est un concept théorique singulier, mais qui est compris comme étant constitué d'un nombre interrelié de dimensions qui « reflètent » des aspects différents de ce même concept (Law et coll., 1998).

Dans la littérature en SI, par exemple, le concept de la confiance a été souvent conceptualisé et mesuré comme un construit multidimensionnel. McKnight et coll. (2002) ont modélisé la confiance dans les contextes organisationnels en fonction de trois facteurs / dimensions : (1) la bienveillance (le degré avec lequel on perçoit que l'autre partie veut notre bien), (2) la compétence (le degré d'aptitudes et de compétences de l'autre partie), et (3) l'intégrité (la perception que l'autre partie adhère à un ensemble de normes). Voir Wright, Campbell, Thatcher et Roberts (2012) pour une discussion plus détaillée des mesures multidimensionnelles.

Étant donné que le concept de qualité peut toucher à plusieurs aspects, Garvin (1987) chercha à rassembler les dimensions de qualité les plus générales qui peuvent s'appliquer dans de nombreux domaines. David A. Garvin, un spécialiste dans le domaine du contrôle de la qualité, a soutenu en 1987 que la qualité peut être utilisée de manière stratégique pour rendre la concurrence plus efficace. Selon l'auteur, une stratégie de qualité appropriée devrait prendre en compte différentes dimensions importantes. Nous les expliquons dans le tableau suivant :

Tableau 2.2 – Dimensions de la qualité selon Garvin (1987)

Dimensions	Explications
Performance	Les différents aspects du fonctionnement d'un produit.
Caractéristiques	Les aspects complémentaires aux aspects de fonctionnement d'un produit.
Fiabilité	Le degré de la confiance dans le fonctionnement du produit durant une longue période de temps.
Conformité	C'est-à-dire dans quelle mesure le produit est conforme aux spécifications préétablies.
Durabilité	La durée durant laquelle un produit peut être utilisé avant que son remplacement devienne nécessaire.
Facilité d'entretien	La rapidité, la courtoisie, la compétence et la facilité de réparation lorsqu'un produit se détériore et doit être retourné pour réparation.
Esthétique	L'impact sur les sens humains (l'apparence, le son, le goût, etc.).
Qualité perçue	La perception de la qualité du produit dans l'esprit du consommateur.

2.5.3 Qualité et Internet

L'utilisation d'Internet pour rentabiliser les affaires a contribué à l'intérêt de plusieurs chercheurs à explorer le concept de qualité dans le contexte de la vente en ligne (Loiacono et coll., 2002, 2007 ; Peterson et coll., 1997). Avec l'augmentation de la concurrence commerciale, les entreprises présentes sur le Web cherchent toujours à améliorer leurs produits et services en se basant sur les perceptions quant à la qualité

des acheteurs provenant de différentes sources (par ex. les feedbacks laissés par les clients sur le site Web du marchand). Les recherches démontrent clairement que l'amélioration de la qualité de service en ligne permet d'augmenter la satisfaction du client et l'intention d'achat, contribue à attirer un nombre plus élevé d'acheteurs sur le site, et peut même contribuer à fidéliser les clients (Ibzan et coll., 2016 ; Shafiee et coll., 2018).

2.5.4 Qualité et commerce électronique

Les services de transaction en ligne ont été mis à la disposition des compagnies commerciales grâce à la technologie de l'information basée sur le Web. Ces services (appelés « *systèmes de commerce électronique* ») jouent à titre d'intermédiaires permettant d'établir des interactions en ligne entre les entreprises et les clients (Evanschitzky et coll., 2004 ; Isaac et Volle, 2014; Meuter et coll., 2000). Pavlou et Gefen (2004) définissent le système de commerce électronique comme étant un système d'information (SI) sur le Web qui offre des services de transaction en ligne pour les acheteurs et les vendeurs. Un fournisseur de services de commerce électronique est une organisation tierce qui utilise des systèmes de commerce électronique pour faciliter les transactions entre acheteurs et vendeurs sur son marché en ligne en collectant, en traitant et en diffusant des informations (voir aussi Sarkar et coll., 1995). Barnes et Vidgen (2002) croient que le défi majeur des organisations de commerce électronique est de comprendre les besoins des clients et d'augmenter leur présence sur le Web. Une organisation avec un site Web ou un « app » difficile à utiliser et à interagir projettera une mauvaise image et affaiblira la position de l'organisation (Connolly et coll., 2018). Il est donc important qu'une organisation soit en mesure d'évaluer la qualité de son offre de commerce électronique telle qu'elle est perçue par ses clients. DeLone et McLean (2003) déclarent qu'il existe trois types d'aspects importants dans les systèmes de commerce électronique : le contenu des sites Web,

l'aspect technique et les services fournis. La qualité de ces aspects (étiquetés d'après la qualité de l'information, la qualité du système et la qualité du service, respectivement) peut avoir une incidence sur les avantages commerciaux enracinés dans l'engagement des clients et leur rétention.

2.5.5 Catégories de la qualité dans le commerce électronique

D'après ce que nous avons relevé dans la littérature précédemment exposée (à partir de la section 2.5.4), nous avons vu que d'importants efforts ont été déployés pour étudier assez exhaustivement le concept de qualité, de l'inclure dans tout ce qui est du Web (commerce électronique, sites Web, agents conversationnels, etc.), ainsi qu'évaluer les différents systèmes d'informations en terme de qualité. Nous en avons conclu qu'il sera utile pour notre étude de discuter des cinq (5) grandes catégories de qualité qui servent à évaluer les services du commerce électronique : la qualité de l'information, la qualité du système, la qualité du service, la qualité de l'interaction, et la qualité du site Web. Ces catégories ont été proposées car la revue de littérature a montré des proximités remarquables entre des dimensions de qualité, ce qui nous a permis de les catégoriser de cette manière. De plus, étant donné que dans notre étude, l'évaluation de la qualité diffère un peu de celle appliquée dans les services de commerce électronique (évaluation de la qualité des chatbots), nous avons considéré qu'il est important pour cette étude d'inclure deux autres catégories : l'apparence visuelle et l'aspect innovant du chatbot. Nous expliquerons le tout en détail dans les sections qui suivent.

2.5.5.1 Qualité de l'information

Parmi les chercheurs qui ont travaillé sur le concept de qualité de l'information, l'on retrouve DeLone et McLean (1992). Ces chercheurs ont créé un modèle qui pourrait servir à mesurer le succès des systèmes d'information. Dans leur modèle, les auteurs

ont inclus la qualité de l'information comme étant l'une des plus importantes dimensions de qualité de systèmes. La qualité de l'information sert à mesurer la sortie du système d'information (niveau sémantique). Selon leur modèle, la qualité de l'information indique le degré auquel l'information est personnalisable, complète, pertinente, sécurités et facile à comprendre. Selon leur modèle, si l'information générée par un système d'information respecte ces critères, les personnes l'utiliseront. Et même lorsqu'ils ont mis à jour leur modèle en 2003, ils n'ont pas mis au rancart leur modèle, ces critères constituent donc l'une des dimensions importantes dans leur modèle (DeLone et McLean, 2003). Dans le tableau 2.3, nous montrons les mesures de succès du commerce électronique utilisées dans le modèle de DeLone et McLean.

Tableau 2.3 – Mesures de succès du commerce électronique du modèle de DeLone et McLean

Dimensions	Métriques
Qualité des systèmes	Adaptabilité
	Disponibilité
	Fiabilité
	Temps de réponse
	Convivialité
Qualité de l'information	Complétude
	Facilité de compréhension
	Personnalisation
	Pertinence
	Sécurité
Qualité du service	Assurance
	Empathie
	Réactivité
Utilisation	Nature de l'utilisation
	Modèles de navigation
	Nombre de visites sur le site
	Nombre de transactions exécutées
Satisfaction de l'utilisateur	Répéter les achats
	Répéter des visites
	Enquêtes auprès des utilisateurs

Avantages nets	Économies de coûts
	Marchés élargis
	Ventes additionnelles incrémentales
	Réduction des coûts de recherche
	Économies de temps

2.5.5.2 Qualité de système

Kuligowska (2015) a touché sur la qualité de service quand elle a développé une mesure pour évaluer la performance des chatbots. Elle a ciblé les assistants virtuels en ligne qui sont capables de converser en polonais, et qui ont été créés par les six grandes compagnies suivantes œuvrant dans le développement de chatbots : Artificial Solutions, Denise Systems, Fido Intelligence, Inteliwise, Online Tools et Stanusch Technologies, et ce, étant donné qu'à l'époque, ces compagnies ont dominé le marché du domaine des chatbots. Kuligowska a aussi visé des chatbots qui comportent des caractéristiques spéciales : ceux qui sont des assistants virtuels commerciaux, ceux qui opèrent dans le secteur B2C, ceux qui atteignent la plus large gamme possible d'utilisateurs, et ceux qui sont probablement les déploiements commerciaux les plus avancés de leurs créateurs. Au début de son étude, Kuligowska (2015) s'est basée sur trois aspects importants : l'évaluation de la performance, la facilité d'usage et la qualité globale d'un agent conversationnel. Après cette étude, dix aspects différents du fonctionnement de chaque assistant virtuel ont été soumis à leur analyse complète. Ils sont montrés dans le tableau 2.4 :

Tableau 2.4 – Les dix différents aspects du fonctionnement de chaque assistant virtuel (Kuligowska, 2015)

Les dix aspects différents du fonctionnement de chaque assistant virtuel
Apparence visuelle
Forme de mise en œuvre sur le site Web
Unité de synthèse vocale
Base de connaissances intégrée
Présentation de connaissances et fonctionnalités supplémentaires
Capacités de conversation et contexte La sensibilité
Traits de personnalité
Options de personnalisation
Réponses d'urgence dans des situations inattendues
Possibilité de classer le chatbot et le site Web par l'utilisateur

2.5.5.3 Qualité de service

Le concept de qualité a été discuté et étudié en profondeur dans le domaine du marketing des services (*services marketing*). Plusieurs chercheurs considèrent la qualité de service comme étant la différence entre les attentes des clients concernant le service attendu et le service perçu. Si les attentes sont supérieures à celles du rendement, la qualité perçue est moins satisfaisante, et par conséquent, l'insatisfaction des clients se produit (par ex. : Shahin, 2006 ; Zeithaml et coll., 1985 ; Parasuraman et coll., 1988).

Dans les années 1980, Zeithaml et coll. (1985) ont visé l'étude de la qualité dans les secteurs de services. Ils ont développé un instrument (baptisé « *SERVQUAL* ») pour mesurer la qualité de service. *SERVQUAL* a été initialement conçu pour mesurer dix aspects de la qualité des services : fiabilité, réactivité, compétence, accès, courtoisie, communication, crédibilité, sécurité, compréhension du client et tangibilité. En se basant sur leur définition de la qualité des services, *SERVQUAL* fut opérationnalisé de manière à mesurer l'écart entre les attentes des clients et leurs expériences. En 1988,

les dix composantes ont été réduites à cinq dimensions dénommées « *R-A-T-E-R* »¹³. La définition de chacune des cinq dimensions est disponible dans le tableau 2.5. Leurs analyses statistiques de leur modèle initial ont permis d'identifier la fiabilité, les tangibles et la réactivité comme étant des dimensions distinctes. Les sept composantes restantes ont été réduites en deux dimensions agrégées : l'assurance et l'empathie (Parasuraman et coll., 1988). Parasuraman et coll. (1988) ont développé un instrument comportant 22 échelles pour mesurer les attentes et les perceptions des clients sur les cinq dimensions (RATER) dans le tableau suivant. Quatre ou cinq items dans le questionnaire sont utilisés pour mesurer chaque dimension.

Tableau 2.5 – Les cinq dimensions de SERVQUAL (Parasuraman et coll., 1988)

Dimensions	Explication
Fiabilité	Capacité à effectuer le service promis de manière fiable et précise.
Tangibles	Apparition d'installations physiques, d'équipement, de personnel et de matériel de communication.
Réactivité	Volonté d'aider les clients et fournir un service rapide.
Assurance	Connaissance et courtoisie des employés et leur capacité à transmettre la confiance.
Empathie	Soins, attention individualisée que l'entreprise offre à ses clients.

SERVQUAL demeure la tentative la plus complète pour conceptualiser et mesurer la qualité du service. Le principal avantage de cet instrument de mesure tient à la capacité des chercheurs à examiner de nombreuses industries de service (comme les soins de santé, les services bancaires, les services financiers et l'éducation) (Shahin, 2006). SERVQUAL a été reconnu comme présentant plusieurs avantages, mais aussi des

¹³ En référence à *Reliability* (fiabilité), *Assurance*, *Tangibles*, *Empathy* (empathie) et *Responsiveness* (réactivité).

limitations. Shahin (2006) a discuté des avantages et des inconvénients du modèle SERVQUAL. Parmi les avantages, on retrouve les suivants :

- Il permet d'évaluer la qualité du service du point de vue du client.
- Il a la capacité de suivre les attentes et les perceptions des clients au fil du temps, ainsi que les divergences entre eux.
- Il permet une comparaison avec les concurrents sur des aspects communs.
- Il a la capacité d'évaluer les attentes et les perceptions des clients internes.

Mais une des limites les plus importantes liées à SERVQUAL, c'est que l'applicabilité uniforme de la mesure dans tous les secteurs de service est difficile. Pour cette raison, des chercheurs dans des domaines comme le commerce électronique et les services de la santé ont mené des recherches afin de développer des mesures de qualité de service mieux contextualisées à certains types de services.

2.5.5.4 Qualité de l'interaction

Conceptualiser la qualité de l'interaction fut considéré comme un défi en raison de sa nature dynamique (Alotaibi et coll., 2011). La qualité d'interaction fait référence à l'expérience d'un client lorsqu'il établit une interaction avec un élément humain d'une organisation de service (Alexandris et coll., 2006 ; Brady et Cronin, 2001). Certains chercheurs affirment que les aspects liés à l'interaction personnelle durant la délivrance de service sont les plus importants pour permettre aux entreprises d'exceller et, potentiellement, de dominer la concurrence (Parasuraman et coll., 1991).

Alotaibi et coll. (2011) ont mené des recherches visant à mesurer la qualité de l'interaction durant la délivrance de services dans l'industrie de l'hôtellerie. Cette recherche a servi à développer et valider une échelle de 24 éléments. Leurs recherches ont été menées en trois étapes pour générer un groupe d'articles d'échelle : (1) une

revue de la littérature (où ils ont présenté leurs premiers modèles de qualité de service), (2) des entrevues qualitatives approfondies, et (3) une enquête quantitative afin de valider leur modèle (Alotaibi et coll., 2011). Le processus de développement de cette échelle est décrit dans les lignes suivantes.

Générer des items à partir de la littérature

Alotaibi et coll. (2011) ont employé les aspects sociaux de la qualité de l'interaction étant donné que de nombreuses recherches ont mis en évidence leur importance dans le processus de prestation des services. Pour cette raison, ils ont proposé la *communication sociale* (entre acheteur et vendeur) comme étant la première dimension pour mesurer la qualité d'interaction. Comme deuxième dimension de qualité d'interaction, Alotaibi et coll. (2011) ont proposé la compétence des employés lors de l'interaction. Ils se sont basés sur les arguments de Hausman (2003) qui a affirmé que les aspects sociaux seraient insuffisants si l'employé du service ne fournit pas les éléments nécessaires de base du service.

Planifier des entrevues semi-structurées

Alotaibi et coll. (2011) ont ensuite mené des entrevues semi-structurées avec un échantillon de participants pour valider et clarifier les concepts qu'ils cherchaient à valider. L'objectif de ces entrevues était d'étudier les dimensions sur lesquels les clients se basent pour évaluer la qualité de l'interaction lors d'une demande de service.

Planifier un sondage quantitatif

Finalement, Alotaibi et coll. (2011) ont élaboré un questionnaire en fonction de leur revue de la littérature et des résultats obtenus durant les entrevues qualitatives. Ce questionnaire comporte trois parties : la première partie mesure la qualité de l'interaction (avec les dimensions identifiées ci-haut), la deuxième partie mesure la

satisfaction de la rencontre, et la troisième partie cherche à mesurer les intentions comportementales des acheteurs. Proposé par Alotaibi et coll. (2011), le tableau 2.6 présente le cadre conceptuel de la qualité de l'interaction dans l'échange de services :

Tableau 2.6 – Le cadre conceptuel de la qualité de l'interaction dans l'échange de services.

Dimensions	Sous-dimensions
Qualité de l'interaction	Communication sociale
	Compétence et assistance liées aux tâches
Conséquences (satisfaction de la rencontre)	Fidélité
	WOM (<i>Word of Mouth</i>)

2.5.5.5 Qualité des sites Web

Avec le développement de la technologie Web, de nombreux chercheurs se sont mis à étudier la qualité des sites Web. Une raison principale pour cet attrait, c'est que les sites Web commerciaux sont devenus une composante essentielle de la croissance rapide du commerce électronique. Loiacono et coll. (2002) ont développé un instrument de qualité pour les sites Web : WebQual. Son développement s'est basé sur une vaste revue de la littérature et des entretiens avec des concepteurs et des visiteurs Web. L'instrument a été raffiné en utilisant deux échantillons successifs, et la validité de l'instrument final a été testée avec un troisième échantillon confirmatoire. Les catégories initiales, représentées comme des dimensions distinctes de la qualité du site Web, apparaissent dans le tableau suivant :

Tableau 2.7 – Les dimensions de la qualité des sites Web (Loiacono et coll., 2002)

Catégories	Dimensions
Facilité d'usage	Facilité de compréhension
	Opérations intuitives
Utilité	Qualité d'information
	Adéquation de l'information à la tâche (<i>Informational Fit-to-Task</i>)
	Interactivité
	Confiance
	Temps de réponse
Divertissement	Attrait visuel
	Innovation
	Flux
Relation complémentaire	Image cohérente
	Intégrité on-line
	Avantage relatif

Loiacono et coll. (2002) ont effectué la procédure de développement de leur instrument en passant par quatre étapes :

- **Étape 1 (définir les dimensions de WebQual)** – Ils ont commencé par 5 dimensions initiales subdivisées en 14 dimensions distinctes. Ils les ont tirées de la revue de littérature en se basant sur leurs définitions (facilité d'usage, utilité, divertissement, relation complémentaire, service client).
- **Étape 2 (développer les items)** – Un ensemble initial de 142 items candidats a été développé sur la base de 13 concepts issus de la revue de la littérature et des études exploratoires. Cette liste d'items a ensuite été affinée sur la base d'une approche utilisée par Davis (1989) dans son prétest de mesures pour le TAM.
- **Étape 3 (raffinement)** – L'instrument a été affiné en examinant sa fiabilité et sa validité discriminante après chacune des deux administrations distinctes. Avec 88 questions dans le questionnaire initial et une règle de base pour

l'analyse factorielle d'au moins cinq fois plus d'observations que de variables à analyser, un nombre d'au moins 440 sujets fut nécessaire. Les données ont été recueillies auprès de 510 étudiants du premier cycle.

- **Étape 4 (sélection et évaluation finale des mesures)** – Pour tester la validité de la mesure de la deuxième version de l'instrument, les auteurs ont procédé à une deuxième série de collecte de données (recueillies auprès de 336 étudiants de premier cycle), et ils ont utilisé deux méthodes pour effectuer les tests de la validité discriminante : l'« *Exploratory Factor Analysis (EFA)* » et la « *Confirmatory Factor Analysis (CFA)* ».

Autres études sur la qualité des sites Web

Il n'y a pas que Loiacono et coll. (2002) qui se sont intéressés à la qualité des sites Web. Plusieurs autres chercheurs ont évolué dans ce domaine, y compris Barnes et Vidgen (2002), ainsi que Wolfinbarger et Gilly (2003). En 2002, Barnes et Vidgen (2002) proposent WebQual, version n° 4 (une version inspirée de l'instrument de Loiacono et coll. [2002]). Cet instrument a été créé pour évaluer la qualité des sites Web d'achat de livres en ligne, et il a été raffiné en utilisant différents sites de commerces en ligne ainsi que des sites de services gouvernementaux. Nous démontrons les dimensions de cette mesure et leurs définitions dans le tableau suivant :

Tableau 2.8 – Échelle WebQual 4.0 (Barnes et Vidgen, 2002)

Dimensions	Items
Convivialité	Facilité pour apprendre à utiliser le site. L'interaction est facile à comprendre. La navigation est facile. Le site est facile à utiliser. Le site est visuellement attirant. Le design est approprié pour le type de site. Le site procure une expérience positive.
Qualité de l'information	L'information est pertinente. L'information est réelle. L'information est actuelle. L'information est facile à comprendre. L'information est précise. Le format utilisé pour présenter l'information est adéquat.
Interaction	Les données personnelles sont confidentielles. L'information est personnalisée. Il est facile de communiquer avec l'entreprise.

Après environ un an, Wolfinbarger et Gilly (2003) ont développé la mesure « *eTailQ* » qui a souligné comme résultat la fiabilité et le design du site comme étant les dimensions les plus importantes pour la qualité des sites Web. Leur échelle inclut quatorze items identifiant quatre dimensions de qualité : design du site Web, service à la clientèle, fiabilité et réalisation, et sécurité / confidentialité. Nous présentons cet instrument dans le tableau suivant.

Tableau 2.9 – Échelle eTailQ (Wolfinbarger et Gilly, 2003)

Dimensions	Items
Design	Navigation. Recherche de l'information. Facilité à conclure les transactions. Clarté de l'organisation et arrangement du site. Personnalisation de la configuration.
Fiabilité et réalisation	Les produits livrés sont les produits commandés. Respect des engagements de temps et de qualité. L'image des produits correspond aux produits commandés.
Sécurité / Confidentialité	Sécurité des paiements. Confidentialité. Le site semble sécurisé.
Service à la clientèle	Assistance en ligne. Réponse immédiate aux clients. Intérêt dans la résolution des problèmes.

2.5.5.6 Anthropomorphisme

Le terme « *Anthropomorphisme* » vient d'une expression grecque composée de deux mots : « *anthropos* » pour « *humain* », et « *morphe* » pour « *forme* ». Le terme décrit la tendance des gens à attribuer des caractéristiques humaines à des artefacts non vivants (Duffy, 2002). Il s'agit d'attribuer des intentions et des émotions à des objets pour les rendre plus familiers, explicables ou prévisibles (Admoni et Scassellati, 2012). De nos jours, l'anthropomorphisme joue un rôle important dans la conception de robots socialement interactifs, et cela se reflète dans leurs formes, leurs comportements et leurs interactions (Fink, 2012).

Nous pensons qu'il est très important d'inclure cette composante interactionnelle de la qualité dans notre étude, car l'intégration des caractéristiques humaines (la convivialité, la personnalisation des communications entre l'utilisateur et le chatbot, etc.) dans la conception des chatbots est devenue de plus en plus demandée. La raison pour cela,

c'est de rendre les chatbots plus similaires au comportement humain, ce qui aide à leur faire prendre des fonctions humaines dans la communauté.

Dans ce qui suit, nous expliquerons l'impact de l'anthropomorphisme au niveau de la forme, du comportement et de l'interaction des robots :

Forme du robot

Fink (2012) considère que ce qui rassemble le design anthropomorphique, qui contient des caractéristiques humaines et la robotique sociale, c'est le fait que l'apparence et la fonction d'un produit influent sur la façon dont les gens perçoivent, interagissent et construisent des relations à long terme avec le produit. Ainsi, l'apparence visuelle d'un robot doit correspondre à ses capacités ainsi qu'aux attentes des utilisateurs.

Comportement et interaction du robot

Fong et coll. (2003) pensent les robots sociaux peuvent utiliser des caractéristiques « *humaines sociales* », par exemple exprimer des émotions, apprendre des modèles d'autres agents, établir des relations sociales, etc. (Fong et coll., 2003). D'autres constatent que toutes les fois où les robots montrent un comportement intentionnel, les gens ont tendance à les percevoir comme des personnages, voire des créatures (Reeves et Nass, 1996 ; Oberman et coll., 2007). Pour cette raison, améliorer l'aspect interpersonnel du comportement du robot devient de plus en plus une nécessité. Selon Craig Besnoy, un consultant à Mindtree, interagir avec des chatbots devrait être comme interagir avec un ami qui nous connaît bien. Cela non seulement permet aux consommateurs d'avoir une conversation en tête-à-tête avec la marque mais ça donne

aussi à celle-ci l'occasion de créer une expérience totalement personnalisée qui est fondée sur les demandes uniques de chaque client¹⁴.

2.5.6 Qualité et chatbots

Au cours des dernières années, les chatbots sont devenus très populaires étant donné l'intérêt grandissant des sociétés de marketing à les utiliser. La raison en est que ce type d'agent conversationnel s'est révélé qualifié pour opérer dans plusieurs domaines humains mieux que l'homme lui-même. Pour cela, les développeurs de chatbots bénéficient désormais d'un soutien important de grandes entreprises de communication (comme Google, Facebook, etc.) (Siddharth et coll., 2015). Cependant, assurer la qualité des chatbots, étant donné leur particularité comparativement à d'autres interfaces, est un peu différent.

Radziwill et Benton (2017) ont mené une étude pour identifier les éléments de la qualité des chatbots et des agents conversationnels. Dans leur étude, ils ont procédé à deux types d'expérimentations : le premier vise à évaluer deux ou plusieurs versions du système conversationnel (évaluer deux chatbots qui existent en même temps) ; le deuxième vise à évaluer un système « *en tant que tel* » et un ou plusieurs systèmes « *à être* » en cours de développement (évaluer un chatbot avant et après les changements).

Radziwill et Benton (2017) ont extrait les éléments de qualité à partir d'un nombre considérable d'articles différents, et ils les ont regroupés en fonction de leur similarité. Ils affirment que ces éléments de qualité peuvent être utilisés comme une liste de

¹⁴ Besnoy, C. (October 20, 2016). Putting personalisation at the heart of the chatbot experience.

Disponible sur : <https://www.itproportal.com/features/putting-personalisation-at-the-heart-of-the-chatbot-experience/>. Dernière visite : 14 Octobre, 2018

contrôle pour une équipe d'implémentation de chatbots afin de s'assurer qu'ils ont résolu les problèmes clés de qualité. Nous les montrons dans le tableau suivant :

Tableau 2.10 – Éléments de qualité (ou reliés à la qualité) pour les chatbots et les agents conversationnels (source Radziwill et Benton, 2017)

Éléments de (ou reliés à) la Qualité des Chatbots	
<i>Efficacité</i>	
Catégories	
Performance	Dégradation gracieuse.
	Robustesse à la manipulation.
	Robustesse à l'entrée inattendue.
	Évite des énoncés inappropriés et capacité à effectuer un contrôle dommage.
	Allocation efficace des fonctions, fourniture des canaux d'escalade appropriés aux humains.
<i>Effectivité</i>	
Catégories	
Fonctionnalité	Synthèse vocale précise.
	Interprète les commandes avec précision.
	Utiliser des degrés appropriés de formalité, registre linguistique.
	Exactitude linguistique des sorties.
	Exécuter les tâches demandées.
	Faciliter les transactions et assurer le suivi des rapports d'état.
	Facilité d'usage générale.
	S'engager dans la résolution de problèmes <i>on-the-fly</i> .
	Contient l'étendue des connaissances, et est flexible dans l'interprétation.
Humanité	Réussite du test de Turing.
	Transparent à l'inspection, divulgue son identité de chatbot.
	Inclure des erreurs pour augmenter le réalisme.
	Convaincante, satisfaisante, interaction naturelle.
	Capable de répondre à des questions spécifiques.
	Capable de maintenir une discussion thématique.

<i>Satisfaction</i>	
Catégorie	
Aspect Affectif	Exprimer des salutations, transmettre la personnalité.
	Donner des repères conversationnels.
	Fournir des informations émotionnelles à travers le ton, l'inflexion et l'expressivité.
	Exprimer la chaleur et l'authenticité.
	Rendre les tâches plus amusantes et intéressantes.
	Divertir et / ou permettre au participant de profiter de l'interaction.
	Lire et répondre aux humeurs d'un participant humain.
Éthique et comportement	Respect, inclusion et préservation de la dignité (liés au choix de l'ensemble d'entraînement).
	Éthique et connaissance culturelle des utilisateurs.
	Protéger et respecter la vie privée.
	Non-déception.
	Sensibilité à la sécurité et aux préoccupations sociales.
	Solvabilité (liée à la qualité perçue).
	Connaissance des tendances et du contexte social.
Accessibilité	Répond aux signaux sociaux ou à l'absence de ceux-ci.
	Peut détecter le sens ou l'intention.
	Répond à différents besoins neurologiques comme le temps de réponse supplémentaire et l'interface de texte.

Par contre, certaines limites peuvent être décelées se rapportant à l'approche qu'ont pris ces auteurs. Premièrement, bien qu'ils aient développé leur liste en se basant sur une revue de la littérature qui touchait de manière prédominante sur l'aspect interactionnel entre l'utilisateur et le chatbot, ils n'ont pas validé empiriquement les catégories qu'ils ont proposées. Deuxièmement, bien que certains des aspects identifiés touchent sur le bon fonctionnement technique du chatbot, aucune de leur catégorie ne prend compte de la qualité de l'information livrée par le chatbot (en d'autres termes, la qualité des recommandations).

2.6 Conclusion au chapitre

Ce chapitre fut consacré à la présentation d'une revue de littérature en exposant différentes recherches antérieures liées à notre sujet d'étude. Dans ce chapitre, nous avons parlé des caractéristiques de qualité des applications de commerce électronique en général et des agents conversationnels en particulier. Nous avons parlé des différents types d'interactions avec les sites Web du commerce, dépendamment de leur complexité de navigation. Nous avons traité du concept d'agent conversationnel et de ses différents types, du point de vue usage et du point de vue stratégie d'implantation, de leurs points communs et de leurs différences.

CHAPITRE III

MODÈLE CONCEPTUEL ET APPROCHE MÉTHODOLOGIQUE

3.1 Introduction au chapitre

Ce chapitre vise à conceptualiser une mesure multidimensionnelle de qualité pour les chatbots. Notre modèle est proposé à partir d'une revue de la littérature (chapitre II). Cette revue de littérature expose des recherches qui ont précédemment développé des échelles pour mesurer la qualité de plusieurs concepts ainsi que la qualité de service dans les environnements commerciaux traditionnels et en ligne. Cela comprend la qualité des systèmes, la qualité des informations fournies, la qualité des interactions, la qualité des services et la qualité des sites Web.

En fait, il existe une vaste littérature sur les perceptions de qualité. Nous avons pris en considération toutes les informations nécessaires. Cependant, notre revue de la littérature a aussi révélé que peu d'études ont été menées sur la façon dont les clients perçoivent la qualité d'un chatbot. Par conséquent, notre développement d'une mesure pour évaluer les perceptions de qualité d'un chatbot est essentiel.

Le présent chapitre expose la méthode utilisée pour mesurer la qualité. Nous avons choisi la méthode proposée par Moore et Benbasat (1991) pour le développement de mesures, laquelle méthode comprend trois étapes : (1) la création d'items, (2) le développement d'une l'échelle, et (3) un test d'instrument. Nous avons présenté nos

dimensions selon différentes catégories. Pour chaque dimension, nous avons proposé une définition et une liste d'items qui ont été inclus dans notre questionnaire. Nous terminerons ce chapitre avec une conclusion et les limites de notre étude.

3.2 Modèle proposé

Notre objectif principal est de proposer et tester empiriquement une mesure multidimensionnelle qui sera utilisée pour mesurer la qualité perçue d'un chatbot. Pour cela, nous avons commencé en identifiant des catégories générales afin de rassembler nos dimensions sous des thèmes semblables. Toutes définies, ces catégories s'énumèrent ainsi : « *Qualité de l'information* », « *Qualité du système* », et la « *Qualité de l'interaction* » entre l'utilisateur et le chatbot. Pour chaque catégorie, nous avons identifié plusieurs sous-dimensions. Nous identifions aussi deux dimensions additionnelles pour la mesure de qualité des chatbots: « *Apparence Visuelle* » et « *Innovation* ». Le modèle proposé apparaît dans la figure 3.1 ci-dessous.

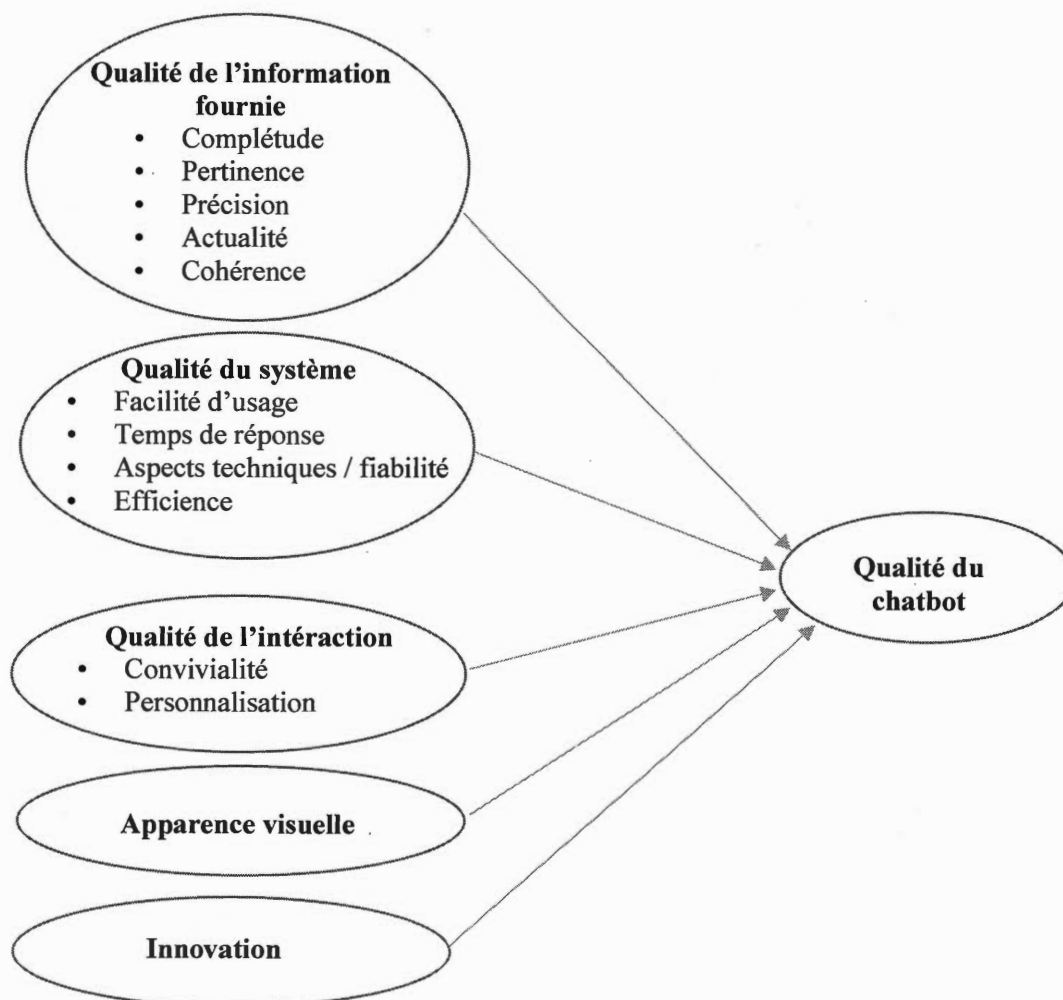


Figure 3.1 – Modèle proposé

Chaque dimension et sous-dimension est abordée dans les sections suivantes. Les items proposés pour les mesurer y sont présentés aussi. En nous basant sur les définitions de ces dimensions et sous-dimensions, nous avons proposé des items qui ont été testés à l'aide d'un questionnaire en ligne. Tous ces items sont de type Likert allant de « 1 » (« *Tout à fait en désaccord* ») jusqu'à « 5 » (« *Tout à fait d'accord* ») – voir questionnaire dans l'annexe F.

3.2.1 Catégorie « *Qualité de l'information* »

Toutes les informations qu'un chatbot offre aux utilisateurs doivent être bien préparées par une équipe associée à l'entreprise qui l'utilise. Comme tout autre système de diffusion d'informations, l'objectif de cela est de garantir la diffusion et l'utilisation rationnelle de l'information (Kottorp et Jäderberg, 2017).

L'information fournie par un chatbot peut varier en fonction des demandes des utilisateurs. Mais selon DeLone et McLean (2003), toute information devrait présenter trois choses : complétude, pertinence et précision. Parce que toute information de bonne qualité qui est placée dans le bon contexte et au bon moment aide à augmenter les profits de l'entreprise, d'autant plus qu'elle peut protéger les personnes contre la tromperie et informer les entreprises de problèmes bien avant qu'ils apparaissent. Pour cela, nous avons conservé les dimensions de qualité utilisées par DeLone et McLean (2003) dans leur modèle de réussite du commerce électronique (la *complétude*, la *pertinence* et la *précision*). Nous croyons que ces dimensions conviennent à notre étude parce que l'information fournie aux clients doit être complète et pertinente, et ce, que la plateforme utilisée soit un site Web, un agent conversationnel ou n'importe quel système d'information. De plus, au moment d'établir le diagnostic de la situation d'un client avec une entreprise, le chatbot doit être précis dans ses interactions avec le client en posant les bonnes questions et dans le bon temps pour identifier le problème.

Selon plusieurs études en qualité de l'information, l'*actualité* de l'information fournie est aussi importante; nous l'avons proposé comme dimension de qualité. Les données publiques des entreprises changent au cours du temps. Les produits et services offerts changent aussi continuellement. Le chatbot devrait être « *conscient* » de ces changements, et l'information qu'il offre aux clients devrait être actuelle, à jour.

Ci-dessous, nous proposons nos quatre (4) premières dimensions pour la qualité des chatbots. Celles-ci correspondent à la *complétude*, *pertinence*, *précision* et *actualité* des information fournies par le chatbots. Nous les présentons en plus de détails dans les sections qui suivent avec les items que l'on a employés dans notre questionnaire pour les mesurer.

3.2.1.1 Dimension « *Complétude des informations* »

Plusieurs chercheurs identifient la complétude comme une caractéristique de la qualité de l'information (par ex., Arazy and Kopak, 2011; Kandari, 2010; Elling et Coll., 2012; Price et coll., 2008). Kandari (2010) définit la complétude des informations comme étant la mesure dans laquelle les informations ne manquent pas et s'avèrent d'une ampleur et d'une profondeur suffisantes pour la tâche à accomplir. Les deux premiers items (Comp1 et Comp2) sont adaptés d'Elling et coll. (2012) :

- *Comp1* – Les réponses fournies sont complètes.
- *Comp2* – Les informations produites par le chatbot sont suffisantes.
- *Comp3* – Ce chatbot répond a toutes les questions d'utilisateur.

3.2.1.2 Dimension « *Pertinence des informations* »

Wang et Strong (1996) définissent la pertinence de l'information comme étant la mesure dans laquelle l'information est applicable et utile pour la tâche à réaliser (Wang et Strong, 1996). L'item Pert2 est adopté de Barnes et Vidgen (2002), et l'item Pert4 est adopté d'Elling et coll. (2012) :

- *Pert1* – Les informations fournies par ce chatbot sont bonnes à savoir.
- *Pert2* – Chaque question / commentaire / réponse était pertinente.
- *Pert3* – Le chatbot a su répondre à mes besoins informationnels.

- *Pert4* – Chaque question / commentaire / réponse était utile.

3.2.1.3 Dimension « *Précision des informations* »

La précision (*accuracy* en Anglais) est la mesure dans laquelle les données décrivent correctement l'objet ou l'événement « *monde réel* » décrit (Askham et coll., 2013). C'est quand l'utilisateur perçoit que les informations sont correctes, stables dans le temps et répétables. Plusieurs études ont identifié ce facteur comme étant une dimension importante de la qualité des informations (par ex., Rieh & Belkin, 1998; Delone & McLean, 1992; Arazy & Kopak, 2011).

- *Prec1* – Chaque question / recommandation provenant du chatbot est précises.
- *Prec2* – Chaque question / recommandation du chatbot était claire pour moi.
- *Prec3* – Certaines communications manquaient de précision (item inversé).
- *Prec4* – Certaines communications me semblaient ambiguës (item inversé).

3.2.1.4 Dimension « *Actualité des informations fournies* »

Appellée *currency*, *recency* ou parfois *timeliness* dans la littérature scientifique publiée en Anglais, l'actualité des informations a été démontré comme étant un aspect important lors de l'évaluation de la qualité d'information par les usagers (Barry, 1994; Schamber, 1991; Rieh & Belkin, 1998; Price, Neiger and Shanks, 2008; Westerman et coll., 2014). Selon l'Institut canadien d'information sur la santé ICIS (2009), l'actualité désigne principalement le caractère courant ou à jour des données au moment de leur diffusion selon l'écart entre la fin de la période de référence à laquelle les données se rapportent et la date à laquelle les données deviennent accessibles aux utilisateurs (ICIS, 2009).

Nous croyons que la même définition s'applique sur les informations fournies par les chatbots ; ceux-ci sont considérés comme des systèmes de diffusion d'informations, et les informations qu'ils fournissent doivent être actuelles et à jour.

- *Actu1* – Les informations fournies par le chatbot semblent récentes.
- *Actu2* – Les informations produites par ce chatbot paraissent être « à jour ».
- *Actu3* – Certaines informations semblaient désuètes (item inversé).

3.2.1.5 Dimension « *Cohérence des informations fournies* »

La cohérence a aussi été identifiée comme une dimension de la qualité des information (Miller, 1996). La cohérence des communications est liée à la cohérence du sujet en discussion. Selon Shen (2013; citant Crystal, 1987), la cohérence signifie qu'il y a une pertinence et une suite logique dans les différents types de concepts et de relations exprimés par un texte. C'est le fait que le sujet abordé par l'utilisateur doit être le même que celui dont parle le chatbot jusqu'à ce que l'utilisateur ait fait le tour du sujet ou décide d'aborder un autre sujet. Miller (1996) décrit cette dimension de la qualité de l'information de la manière suivante : « *(C)oherece is how well the information hangs together and is consistent with itself* » (p. 80). Ainsi, les questions et les réponses du chatbot concernant un même sujet doivent être cohérentes et liées.

- *Cohe1* – Chaque communication était cohérente avec toutes celles qui l'a précédaient.
- *Cohe2* – Les questions posées avaient une suite logique.
- *Cohe3* – Les commentaires du chatbot étaient parfois contradictoires (item inversé).
- *Cohe4* – Certaines communications semblaient etre incohérentes (item inversé).
- *Cohe5* – Les communications sautaient d'un thème à un autre (item inversé).

3.2.2 Catégorie « *Qualité du système* »

Cette catégorie porte sur la qualité de fonctionnement du chatbot et la technologie sur laquelle il se base pour réaliser la performance voulue.

DeLone et McLean (2003), dans leur étude sur la qualité des systèmes, ont identifié le *temps de réponse* et la *facilité d'usage* (parfois appelée « *convivialité* » dans la littérature) comme faisant partie de la qualité nécessaire pour un système de haute qualité. Ces critères sont également applicables aux chatbots. Selon Kuligowska (2015), la *facilité d'usage* et le *temps de réponse* jouent un rôle très important, et ce, parce que la raison fondamentale de favoriser la prise d'informations d'un chatbot plutôt que de la prendre d'un site Web, c'est que l'interaction avec un chatbot en langage naturel va tout simplement faciliter la procédure, et la recherche des informations ne va pas prendre beaucoup du temps (comme la navigation d'un site Web). DeLone et McLean (2003) proposent aussi la *fiabilité* comme facteur de qualité pour les systèmes d'information parce qu'elle est liée principalement aux aspects techniques de leur fonctionnement. Nous voyons que cela s'applique aussi aux chatbots, car ces derniers sont aussi considérés en tant que systèmes qui fournissent des informations.

À ces critères, nous ajoutons « *Aspects techniques / Fiabilité* ». La *fiabilité* est souvent identifiée comme la dimension la plus importante dans la perception de qualité dans les études utilisant la mesure SERVQUAL (Zeithaml et coll., 1985; Parasuraman et coll., 1988).

Finalement, nous proposons aussi que la dimension que nous identifions comme l'*Efficiency* est importante dans les perceptions de qualité chez l'utilisateur des

chatbots. Chacune de ces dimensions est traitée plus en détails dans les paragraphes qui suivent.

3.2.2.1 Dimension « *Facilité d'usage* »

La *facilité d'usage* est définie comme étant le degré auquel une personne croit que l'utilisation d'un système particulier s'effectue sans effort (Loiacono et coll., 2002). La *facilité d'usage* (*ease of use* en Anglais) provient du modèle TAM développé par Davis et coll. (1989). Ces auteurs définissent la *facilité d'usage* comme « *the degree to which the prospective user expects the target system to be free of effort* » (p. 985). Nous croyons que cette dimension pourrait être très importante pour évaluer la qualité des chatbots, car ils sont beaucoup plus faciles à utiliser que d'autres systèmes d'information. L'utilisateur n'a pas besoin de naviguer pour trouver ce qu'il veut, il doit juste poser des questions, et le chatbot répond. Les items *Faci2* et *Faci3* sont adaptés directement de Loiacono et coll. (2002) :

- *Faci1* – L'interface est intuitive.
- *Faci2* – Ce chatbot est facile à utiliser.
- *Faci3* – Je me suis vite habitué à utiliser ce chatbot.

3.2.2.2 Dimension « *Temps de réponse* »

Cette dimension est appelée *response time* dans la littérature scientifique en Anglais. Nous avons trouvé plusieurs définitions pour cette dimension, car elle a été utilisée dans plusieurs domaines. Nous avons adopté celle de Loiacono et coll. (2001) qui l'ont défini comme étant le temps écoulé pour obtenir une réponse après une demande ou une interaction avec un système (Loiacono et coll., 2002 voir aussi Petter et coll., 2008 et Al-Mamary et coll. 2014). Donc, c'est le temps d'attente entre les entrées de l'utilisateur et la réponse du chatbot à ces entrées.

- *Temp1* – Ce chatbot est rapide.
- *Temp2* – Le temps que ca prend au chatbot a répondre est raisonnable.
- *Temp3* – Le temps de réponse est acceptable.

3.2.2.3 Dimension « *Aspects techniques / Fiabilité* »

Parmi les dimensions identifiées dans SERVQUAL, la fiabilité apparaît systématiquement comme étant la dimension la plus importante dans la perception des clients quant à la qualité de service. Lorsqu'il s'agit de chatbots ou d'autres technologies informatiques, la fiabilité est étroitement liée à la qualité technique d'un système (Guesmia et coll., 2018). Shahin (2006) définit la fiabilité comme étant la capacité d'exécuter le service promis de manière fiable et précise. Pour les chatbots, la fiabilité est définie comme étant le degré de bon fonctionnement et l'absence de problèmes techniques qui pourraient autrement interférer avec la capacité du chatbot à répondre aux questions des utilisateurs (Guesmia et coll., 2018).

- *Fiab1* – Ce chatbot fonctionne bien.
- *Fiab2* – Ce chatbot ne m'a pas causé de frustration.
- *Fiab3* – Ce chatbot fonctionne sans problèmes.

3.2.2.4 Dimension « *Efficiencia* »

Selon Ayah et Coll. (2018), il est souhaitable d'avoir un chatbot capable de récupérer des réponses à toutes les questions, non pas seulement de manière rapide mais aussi efficiente. Erik et coll. (2000) définissent l'efficiencia comme la relation entre (1) l'exactitude et l'exhaustivité avec lesquelles les utilisateurs atteignent certains objectifs et (2) les ressources dépensées pour les atteindre. Selon Goh (2013) l'efficiencia consiste à travailler de la meilleure façon possible avec le minimum de perte de temps et d'effort.

- *Effi1* – Obtenir une réponse a pris peu d'efforts.
- *Effi2* – J'ai obtenu les réponses que je cherchais sans trop d'échanges.
- *Effi3* – Mes interactions avec le chatbot ont été efficaces.
- *Effi4* – Le montant d'information échangé entre moi et le chatbot était adéquat.

3.2.3 Catégorie « *Qualité de l'interaction* »

Nous proposons la catégorie « *Qualité d'interaction* » et la définissons comme étant la catégorie qui rassemble toute dimension qui touche les aspects interactifs et relationnels entre l'utilisateur et le chatbot. Kuligowska (2015) croit que les chatbots doivent démontrer la capacité de riches interactions sociales à travers une conversation naturelle, une expressivité flexible et, pour ceux qui de nature anthropomorphique, un comportement gestuel convergent. Nous proposons la convivialité et la personnalisation comme étant des facteurs importants pour assurer la qualité d'interaction d'un chatbot.

3.2.3.1 Dimension « *Convivialité* »

Le chatbot peut afficher des traits de caractère comparables à ceux des êtres humains (politesse, empathie, etc.) ; par conséquent, son interaction avec l'utilisateur doit être amicale et respectueuse. Tout cela fait partie de ce que nous appelons la convivialité.

- *Conv1* – Ce chatbot interagissait avec moi de manière respectueuse.
- *Conv2* – Ce chatbot interagit avec l'utilisateur avec politesse.
- *Conv3* – Mon interaction avec ce chatbot a été amicale.
- *Conv4* – Mes interactions avec ce chatbot m'ont parfois fait sourire.
- *Conv5* – Mes communications avec le chatbot étaient conviviales.

3.2.3.2 Dimension « *Communication personnalisée* »

La personnalisation est définie par Kramer et coll. (2000) comme étant une boîte à outils de technologies et de fonctionnalités d'application utilisées dans la conception d'une expérience d'utilisateur final. Dans le domaine des chatbots, cela peut être vu dans la manière dont le chatbot interagit avec l'utilisateur et répond de manière personnalisé à ses besoins.

- *Comm1* – Je sentais comme une communication personnelle reçue de ce chatbot.
- *Comm2* – Les informations fournies par le chatbot me paraissaient adaptées à mes besoins spécifiques.
- *Comm3* – Le chatbot a ajusté ses réponses et recommandations à mes préférences.

3.2.4 Autres Dimensions Reliées à la Qualité de Chatbots

3.2.4.1 Dimension « *Innovation* »

L'innovation a été adaptée de travaux sur la qualité du site Web (Loiacono et coll., 2002). Elle a été définie comme étant la créativité et l'unicité de la conception du site Web. La même définition s'applique aux chatbots. Pour les chatbots, l'innovation est le degré auquel un utilisateur perçoit le chatbot comme étant créatif et unique (Guesmia et coll., 2018).

- *Inno1* – Ce chatbot est innovant.
- *Inno2* – Ce chatbot est unique.
- *Inno3* – Ce chatbot est technologiquement avancé.

3.2.4.2 Dimension « Apparence visuelle »

« *Un livre est souvent jugé par sa couverture* ». L'attrait visuel a été démontré comme un facteur important influençant les attitudes et les comportements des consommateurs (Guesmia et coll., 2018). Selon Kuligowska (2015), la visualisation d'un chatbot qui ressemble à une personne vivante augmente l'implication de l'utilisateur et sa volonté d'entamer une conversation. Dans le domaine de la qualité des sites Web, Loiacono et coll. (2007) ont défini l'apparence visuelle comme étant le degré auquel l'esthétisme du site Web attire l'utilisateur. Nous adoptons cette définition aux chatbots.

- *Appa1* – L'apparence du chatbot est agréable.
- *Appa2* – L'interface est attirante.
- *Appa3* – Coté visuel, ce chatbot est plaisant.

Dans le tableau suivant, nous rassemblons toutes les catégories, les dimensions et les items du modèle proposé :

Tableau 3.1 – Catégories, dimensions et items du modèle proposé

Catégories	Dimensions	Items
Qualité de l'information	Complétude des informations	Comp1 – Les réponses fournies sont complètes.
		Comp2 – Les informations produites par le chatbot sont suffisantes.
		Comp3 – Ce chatbot répond a toutes les questions d'utilisateur.
	Pertinence des informations	Pert1 – Les informations fournies par ce chatbot sont bonnes à savoir.
		Pert2 – Chaque question / commentaire / réponse était pertinente.
		Pert3 – Le chatbot a su répondre à mes besoins informationnels.
		Pert4 – Chaque question / commentaire / réponse était utile.

Catégories	Dimensions	Items
	Précision	Prec1 – Chaque question / recommandation provenant du chatbot est précises.
		Prec2 – Chaque question / recommandation du chatbot était claire pour moi.
		Prec3 – Certaines communications manquaient de précision (item inversé).
		Prec4 – Certaines communications me semblaient ambiguës (item inversé).
	Actualité des informations fournies	Actu1 – Les informations fournies par le chatbot semblent récentes.
		Actu2 – Les informations produites par ce chatbot paraissent être « à jour ».
		Actu3 – Certaines informations semblaient désuètes (item inversé).
	Cohérence des communications	Cohe1 – Chaque communication était cohérente avec toutes celles qui l'a précédaient.
		Cohe2 – Les questions posées avaient une suite logique.
		Cohe3 – Les commentaires du chatbot étaient parfois contradictoires (item inversé).
		Cohe4 – Certaines communications semblaient être incohérentes (item inversé).
		Cohe5 – Les communications sautaient d'un thème à un autre (item inversé).
Qualité du système	Facilité d'usage	Faci1 – L'interface est intuitive.
		Faci2 – Ce chatbot est facile à utiliser.
		Faci3 – je me suis vite habitué à utiliser ce chatbot.
	Temps de réponse	Temp1 – Ce chatbot est rapide.
		Temp2 – Le temps que ça prend au chatbot à répondre est raisonnable.
		Temp3 – Le temps de réponse est acceptable.
	Aspects techniques / Fiabilité	Fiab1 – Ce chatbot fonctionne bien.
		Fiab2 – Ce chatbot ne m'a pas causé de frustration.
		Fiab3 – Ce chatbot fonctionne sans problèmes.

Catégories	Dimensions	Items
Qualité de l'interaction	Efficience	Effi1 – Obtenir une réponse a pris peu d'efforts.
		Effi2 – J'ai obtenu les réponses que je cherchais sans trop d'échanges.
		Effi3 – Mes interactions avec le chatbot ont été efficaces.
		Effi4 – Le montant d'information échangé entre moi et le chatbot était adéquat.
	Convivialité	Conv1 – Ce chatbot interagissait avec moi de manière respectueuse.
		Conv2 – Ce chatbot interagit avec l'utilisateur avec politesse.
		Conv3 – Mon interaction avec ce chatbot a été amicale.
		Conv4 – Mes interactions avec ce chatbot m'ont parfois fait sourire.
		Conv5 – Mes communications avec le chatbot étaient conviviales.
	Communication personnalisée	Comm1 – Je sentais comme une communication personnelle reçue de ce chatbot.
		Comm2 – Les informations fournies par le chatbot me paraissaient adaptées à mes besoins spécifiques.
		Comm3 – Le chatbot a ajusté ses réponses et recommandations à mes préférences.
Autres Dimensions Reliées à la Qualité de Chatbots	Innovation	Inno1 – Ce chatbot est innovant.
		Inno2 – Ce chatbot est unique.
		Inno3 – Ce chatbot est technologiquement avancé.
	Apparence visuelle	Appa1 – L'apparence du chatbot est agréable.
		Appa2 – L'interface est attirante.
		Appa3 – Côté visuel, ce chatbot est plaisant.

3.3 Hypothèses

À partir du modèle proposé, nous déterminons dans ce qui suit les hypothèses soulignées :

- H1 – La qualité de l'information produite par le chatbot est positivement et significativement reliée à la qualité perçue du chatbot.
- H2 – La qualité de système (du chatbot) est positivement et significativement reliée à la qualité perçue du chatbot.
- H3 – La qualité des interactions entre l'utilisateur et le chatbot est positivement et significativement reliée à la qualité perçue du chatbot.
- H4 – L'apparence visuelle du chatbot est positivement et significativement reliée à la qualité perçue du chatbot.
- H5 – L'innovation perçue du chatbot est positivement et significativement reliée à la qualité perçue du chatbot.

3.4 Validation et Fiabilité

3.4.1 Validation du contenu - Purification des items préliminaires du questionnaire

La validation du contenu est un processus où le chercheur détermine si les éléments d'un instrument d'évaluation sont pertinents et représentatifs de la variable qu'on cherche à mesurer. Pour établir la validité du contenu, nous nous sommes basés sur une méthode utilisée Moore et Benbasat (1991) et Tomiuk (2005) mieux connus sous le nom de tri par carte.

La validation des items et dimensions vise à identifier le degré de pertinence et de représentation des items d'un instrument de mesure (Moore et Benbasat, 1991). Avec une application de la méthode du tri par carte (« *Card sorting* »), nous avons testé si

les items que nous avons proposés correspondaient aux dimensions qu'on cherchaient à mesurer, et pour la réaliser, nous avons invité quatre juges (nommément deux professeurs et deux (2) étudiants de deuxième cycle de l'UQAM qui avaient déjà employé cette procédure) pour établir si l'étendue de nos items de départ représentait bien nos dimensions.

Pour effectuer les tests, nous avons transcrit chaque item proposé sur une carte en papier. Ensuite, nous avons écrit le nom de chaque dimension (ou sous-dimension) et sa définition sur une enveloppe. Une enveloppe supplémentaire intitulée "Incertain" a également été ajoutée ce qui a permis d'y mettre les items dont les juges n'étaient pas certain. Chaque juge a été invité à trier les items et à les mettre dans l'enveloppe correspondante à la dimension ou sous-dimension jugée pertinente. Si au moins trois sur quatre (3/4 ou 75%) des juges associaient les items aux bonnes enveloppes (c.-à-d. dimensions ou sous-dimensions proposées pour notre étude), nous conservions ces items dans le questionnaire final.

3.4.2 Validation du modèle

Afin de purifier notre échelle de qualité multidimensionnelle et de valider nos dimensions proposées, nous avons utilisé une méthode appelée « *Partial Least Square Structural Equation Modeling* » (PLS-SEM), laquelle détermine la relation statistique entre les variables indépendantes (nos dimensions et sous-dimensions de la qualité du chatbot) et la variable dépendante (la qualité globale d'un chatbot).

La régression « *Partial Least Square* » (PLS-SEM) permet d'estimer des modèles de relation de cause à effet complexes en utilisant des variables latentes. Nous avons choisi le SEM-PLS parce qu'il fonctionne bien lorsqu'il utilise des tailles plus petites et des données non normales (Hair et coll., 2011).

Pour le développement de nouvelles mesures, la vérification de la validité convergente devient importante. La validité convergente est le degré dans lequel une mesure est fortement corrélée avec d'autres mesures conçues pour mesurer le même concept (Churchill, 1979). Plus simplement, c'est une vérification que la mesure développer mesure ce que l'on pense qu'elle devrait mesurer.

Afin de tester la validité convergente de notre mesure de qualité des chatbots, nous avons inclus une échelle supplémentaire dans notre questionnaire pour mesurer la qualité globale perçue des chatbots (une mesure unidimensionnelle). Nous proposons les items suivants pour cette mesure de *qualité globale du chatbot* :

- *Qual1* – Ce chatbot est de bonne qualité.
- *Qual2* – Ce chatbot est bien conçu.
- *Qual3* – Je crois que ce chatbot a besoin d'être amélioré (item inversé).

Lors des analyses statistiques effectuées, nous avons évalué la corrélation entre la mesure globale de qualité du chatbot et la mesure multidimensionnelle que l'on a développé dans ce travail. Ces analyses sont présentées dans le chapitre suivant.

3.4.3 Fiabilité

Pour l'analyse de fiabilité de la cohérence interne du construit, nous avons utilisé l'indice alpha de Cronbach. D'après Pedhazur et Schmelkin (1991), le seuil minimum requis est de 0,70. Par contre, pour de nouvelles échelles, on peut être tolérant avec de bas coefficients comme 0,50 et 0,60. La cohérence de l'échelle se base sur la corrélation des éléments entre eux, et sur la corrélation entre chaque élément et l'échelle totale. Nous avons pris 0,70 comme seuil pour la validation de la fiabilité de nos variables.

3.5 Approche pour la collecte de données

3.5.1 Outil de collecte de données

À partir de la méthode de validation que nous avons suivie, nous avons développé un questionnaire en ligne à l'aide de l'outil « *Google Forms* ». Le questionnaire est composé de 52 questions, dont 46 ont été utilisées pour mesurer le construit, les trois autres questions étant de type démographique. Il faut préciser que le questionnaire est de type fermé, et que nous avons utilisé l'échelle de mesure de Likert à cinq points allant de « 1 » (« *Tout à fait en désaccord* ») à « 5 » (« *Tout à fait d'accord* »).

Juste avant de répondre au questionnaire, nous avons mis à la disposition des participants une liste de quatre (4) chatbots¹⁵ (voir le tableau 3.2) pour les faire interagir avec eux durant environ 15 minutes. Cela avait pour objectif de montrer aux participants comment ce type d'agents conversationnels fonctionne afin qu'ils puissent avoir au moins une idée sur les chatbots et ainsi répondre ensuite aux questions. Nous nous sommes assuré que ces quatre chatbots avaient déjà été mis en fonction (jusqu'au moment de la soumission de cette étude). Dans le questionnaire, nous avons demandé à chacun des participants de répondre aux questions en pensant à l'un des chatbots proposés qu'il a le mieux ou le moins appréciés. Cela visait à augmenter la variabilité des réponses aux items de notre questionnaire.

¹⁵ Ces quatre Chatbots ont été choisis parmi une liste initiale de 27 Chatbots en ligne. Chaque chatbot a été visité plusieurs fois par le chercheur principale ainsi que son directeur de mémoire. Certains de ces chatbots semblaient être des « versions bêta » ou en développement, d'autres ne fonctionnaient pas tout le temps et semblaient souffrir de problèmes techniques majeurs et récurrents qui possiblement ne permettraient pas aux répondants dans notre étude d'interagir avec le chatbot de manière continue et suffisante pour permettre la formation d'une opinion sur les diverses dimensions identifiées dans notre questionnaire. D'autres ont été exclus car ni des versions françaises ni anglaises étaient disponibles pour ces chatbots. En somme, la sélection finale était basée sur plusieurs facteurs qui incluaient principalement le nombre et la fiabilité des options disponibles et fonctionnelles dans chaque Chatbot.

Tableau 3.2 – Les quatre chatbots utilisés lors de la collecte de données

Chatbot	Adresse sur le Web
CPF	https://www.moncompteactivite.gouv.fr/cpa-public/
Ask Mona	https://www.messenger.com/t/askmonaparis
Your.MD	https://webapp.your.md/coach
Veronica	https://www.inbenta.com/en/blog/

3.5.2 Administration du questionnaire

Nous avons commencé les démarches de la collecte de données en contactant par courriel les enseignants du cours de MET-1330 pour leur demander s'il leur était possible de nous aider. Les participants étaient des étudiants de l'UQAM inscrits au premier cycle dans les classes du cours MET-1330¹⁶ donné durant la session d'hiver 2018. Aucune restriction n'a été stipulée sur l'âge, le sexe ou l'ethnie. Après avoir obtenu l'accord des enseignants, nous nous sommes présenté au début du cours pour expliquer aux étudiants à quoi servira notre étude. Puis, nous leur avons demandé d'interagir durant 15 minutes avec les chatbots que nous leur avons proposés, pour ensuite répondre au questionnaire en ligne. Le questionnaire a été conçu à l'aide de « *Google Forms* » (un outil de sondage développé par la compagnie Google, et qui permet de créer des questionnaires et les distribuer en ligne).

3.5.3 Estimation de la grandeur de l'échantillon requis

Étant donné que PLS-SEM est encore une méthode d'analyse statistique émergente, on retrouve plusieurs chercheurs qui suggèrent des tailles minimales d'échantillonnage

¹⁶ Cours d'introduction en technologie d'information (TI). Ce cours fait partie des choix de cours dans plusieurs programmes qui incluent le baccalauréat ainsi que plusieurs certificats en administrations/gestion.

différentes. Selon Hair et Coll. (2013), on peut utiliser la « *règle des dix fois* » (« *10 times rule* ») pour estimer la taille minimale de l'échantillon. Selon cette technique, la taille de l'échantillon ne doit pas être inférieure à 10 fois le plus grand nombre de *paths* dirigés vers une variable latente (construit qui reçoit le plus grand nombre de flèches) dans le modèle structurel. Similairement, Marcoulides & Saunders (2006) recommandent une grandeur d'échantillon basée sur le plus grand nombre de *paths* dans le modèle. Le tableau 3.3 ci-dessous montre la grandeur d'échantillonnage requise étant donné le nombre de *paths*.

Tableau 3.3 – Estimation de la taille de l'échantillon selon Marcoulides & Saunders (2006) – Source : Wong (2013).

Échantillon minimum (n=)	Plus grand nombre de flèches pointant à une variable latente du modèle
52	2
59	3
65	4
70	5
75	6
80	7
84	8
88	9
91	10

D'autres chercheurs recommandent une estimation requérant plusieurs paramètres. Il existe des programmes offerts gratuitement en ligne qui nous permettent d'estimer facilement la grandeur de l'échantillon PLS-SEM même si les calculs peuvent être complexes. Un programme très populaire est G*Power (<http://www.gpower.hhu.de>).

Cohen (1992) recommande utiliser les paramètres suivants : un α (probabilité d'erreur) = 0,05, une puissance de test = 0,80 avec une taille médiane de l'effet (f^2) = 0,15.

3.6 Analyse des données

L'analyse des données a été menée à l'aide du logiciel SmartPLS (version n° 3). Nous avons transféré les données de l'outil « *Google Forms* » à un fichier Excel, et les données furent ensuite traitées en utilisant SmartPLS. Nous avons utilisé l'alpha de Cronbach pour analyser la fiabilité de notre mesure, et l'analyse de la méthode PLS pour vérifier si des relations significatives existaient entre les dimensions et la qualité.

3.7 Conclusion au chapitre

Dans ce chapitre, nous avons mis la lumière sur le modèle que nous avons proposé comme mesure multidimensionnelle de qualité des chatbots. Nous avons expliqué les démarches ayant conduit au choix des items et des dimensions de qualité, ainsi qu'aux deux méthodes de validation : la validation préliminaire du regroupement des items et des dimensions, puis l'analyse confirmatoire et la validation de la version finale du modèle. Nous avons aussi discuté la façon dont nous avons distribué le questionnaire qui fut à la base de la réalisation de la deuxième méthode de validation, la population choisie pour répondre au questionnaire et la méthode utilisée pour analyser les réponses. Nous avons aussi montré les outils utilisés pour la création et la distribution du questionnaire et l'analyse de données.

CHAPITRE IV

ANALYSE ET PRÉSENTATION DES RÉSULTATS

4.1 Introduction au chapitre

Dans ce chapitre, nous présenterons les résultats de l'analyse que nous avons effectuée pour valider notre mesure de qualité pour les chatbots. Tout d'abord, nous commencerons par présenter à nouveau les hypothèses. Ensuite, nous discuterons du processus de validation connu sous le nom de « *Card Sorting* ». Ce processus génère un questionnaire qui contient des items validés par des experts que nous avons employés pour élaborer cette méthode. Ensuite, nous exposerons les résultats de l'analyse de la méthode statistique PLS pour déterminer la relation et le degré d'importance de chaque dimension de qualité par rapport à la qualité globale du chatbot. Nous concluons avec un sommaire de notre analyse. Une discussion plus approfondie et une conclusion générale seront présentées dans le chapitre suivant.

4.2 Hypothèses

Selon le modèle proposé, nous avons déterminé les hypothèses suivantes :

- H1 – La qualité de l'information produite par le chatbot est positivement et significativement reliée à la qualité perçue du chatbot.

- H2 – La qualité de système (du chatbot) est positivement et significativement reliée à la qualité perçue du chatbot.
- H3 – La qualité des interactions entre l'utilisateur et le chatbot est positivement et significativement reliée à la qualité perçue du chatbot.
- H4 – L'apparence visuelle du chatbot est positivement et significativement reliée à la qualité perçue du chatbot.
- H5 – L'innovation perçue du chatbot est positivement et significativement reliée à la qualité perçue du chatbot.

4.3 Profil démographique

D'après les réponses que nous avons reçues de la part des répondants ($n = 87$), nous avons pu identifier le profil démographique de l'ensemble de l'échantillon. Nous le décrivons dans le tableau suivant :

Tableau 4.1 – Profil démographique de l'échantillon

Variable	Catégorie	Pourcentage (%)
Sexe	Hommes	46 %
	Femmes	54 %
Âge	(18-22)	32 %
	(23-27)	43 %
	(28-32)	10 %
	(33-37)	8 %
	(38-42)	6 %
	(43-47)	1 %
	(48-52)	0 %
	(53 et +)	0 %
Éducation	Secondaire	12 %
	Cégep	70 %
	Baccalauréat	8 %
	Maîtrise	0 %
	Doctorat	0 %
	Autres diplômes universitaires	10 %

4.4 Validation du modèle proposé

Il faut se rappeler que l'objectif de ce travail est le développement d'une échelle multidimensionnelle qui peut être utilisée d'une part comme une mesure de qualité pour mesurer la perception de la qualité d'un nouveau robot, et d'autre part comme un outil de diagnostic pour évaluer la qualité des chatbots existants. Nous avons utilisé des dimensions que nous avons trouvées dans la littérature et d'autres qui ont émergé durant notre revue de la littérature. Ces dimensions ont été rassemblées par catégorie selon le thème.

4.4.1 Raffinement du questionnaire à l'aide des juges

La validation des items du questionnaire en utilisant une application de la méthode « *tri par carte* » (« *card sorting* » - voir Moore et Benbasat, 1991) a établi l'étendue de nos items de départ. Nous sommes restés avec les items présentés dans ce tableau-ci :

Tableau 4.2 – Dimensions après validation des experts

Catégories	Dimensions
Qualité de l'information	Complétude des informations (Comp1, Comp2, Comp3)
	Pertinence des informations (Pert1, Pert2, Pert3, Pert4)
	Précision (Prec1, Prec2, Prec3, Prec4)
	Actualité des informations fournies (Actu1, Actu2, Actu3)
	Cohérence des communications (Cohe1, Cohe2, Cohe3, Cohe4, Cohe5)
Qualité du système	Facilité d'usage (Faci1, Faci2, Faci3)
	Temps de réponse (Temp1, Temp2, Temp3)
	Aspects techniques / Fiabilité (Fiab1, Fiab2, Fiab3)
	Efficience (Effi1, Effi2, Effi3, Effi4)
Qualité de l'interaction	Convivialité (Conv1, Conv2, Conv3, Conv4, Conv5)
	Communication personnalisée (Comm1, Comm2, Comm3, Comm4)
Autres Dimensions Reliées à la Qualité de Chatbots	Innovation (Inno1, Inno2, Inno3)
	Apparence visuelle (Appa1, Appa2, Appa3)

4.4.2 Raffinement du modèle en utilisant la méthode PLS-SEM

Nous représentons les modèles PLS-SEM en utilisant deux types de variables. Ces deux types de variables sont les « *variables observées* » et les « *variables latentes* ». Les

variables observées sont les items inclus dans le questionnaire. Ces variables sont aussi appelées des « *indicateurs* ».

On utilise les variables « latentes » pour représenter des concepts abstraits qui ne sont pas directement observables ou mesurables. Par exemple, la « *facilité d'usage* » n'est pas directement mesurable et reflète plutôt une perception chez l'utilisateur d'un système d'information. La partie d'un modèle structurel décrivant les relations entre les variables latentes est appelée « *modèle interne* », et les relations stipulées entre les variables latentes sont appelées « *paths* ».

Au lieu d'être directement observables, les variables latentes sont déduites d'autres variables qui pour leur part sont observables et mesurables. Le « *modèle externe* » est alors la partie du modèle qui représente les relations entre les variables latentes et leurs indicateurs (variables observables et mesurables). Les relations stipulées entre une variable latente et ses indicateurs sont appelées des "*loadings*".

Des logiciels permettant de mener des analyses statistiques (comme SmartPLS, LISREL, AMOS, etc.) permettent aux chercheurs de tester un modèle en calculant la partie interne et la partie externe du modèle. On réfère à ces deux parties du modèle global en utilisant respectivement les expressions « *modèle structurel* » et « *modèle de mesure* ». Ces logiciels établissent des estimations de paramètres pour les deux parties du modèle qui sont nécessaires à analyser. Le chercheur utilise ces paramètres pour affiner le modèle en se basant sur des seuils précis établis par la communauté scientifique (par ex., Kock, 2016). L'analyse et l'affinement du modèle sont effectués en utilisant plusieurs critères. Ceux-ci sont présentés dans la section 4.5.

4.4.3 Spécifications Réflectives et Formatives

PLS-SEM nous permet de modéliser des variables latentes à l'aide de conceptualisations formative ou reflective (voir Figure 4.1). L'hypothèse derrière la conceptualisation réflexive est que tout changement dans le variable latente (construit) est supposé causer des modifications dans les variables observables (les indicateurs ou Xs). Les modélisations reflectives sont démontrés avec des flèches quittant le construit qu'on cherche à mesurer. Dans la conceptualisation formative, l'hypothèse qui soutient ces modèles indique que ce sont les éléments mesurés (les Xs) qui entraînent des changements dans le construit (la variable latente). Les flèches sont donc représentées allant des items vers le construit (variable latente).

Dans les modèles reflectifs, les indicateurs sont hautement corrélés ensemble et dans les modèles formatifs, les indicateurs ne le sont pas. Leurs effets sur la variable latente (construit) est plutôt considérés comme étant additifs les uns relativement aux autres (Vlachos & Aristeidis, 2009).

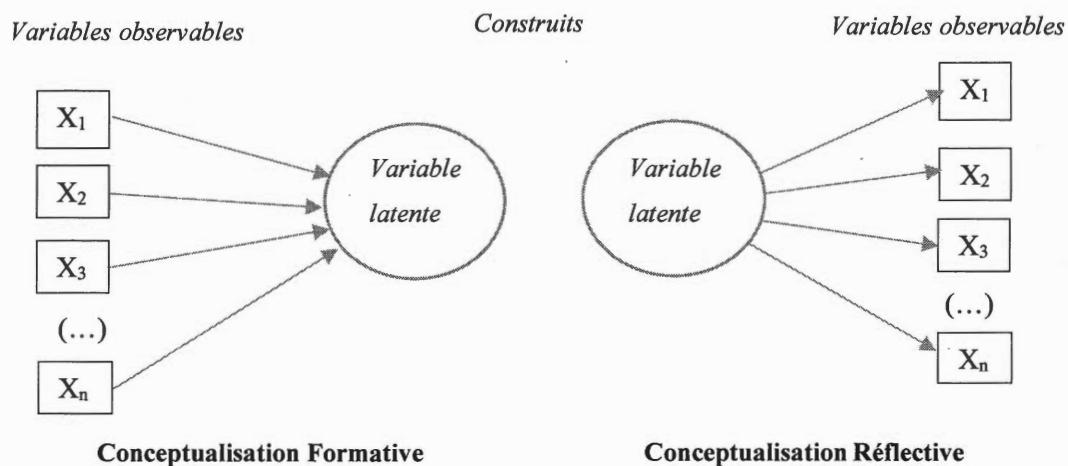


Figure 4.1 – Modélisations de Modèles Formatifs et Réflexifs

Des conceptualisations formatives et réflexives peuvent être combinés ensemble pour former des modèles complexes où l'on retrouve plusieurs construits. La qualité du système est un exemple d'un construit de plus haut niveau qui peut être modélisé en combinant des conceptualisations réflexives et formative. Dans la figure 4.2, la qualité du système peut être conceptualisée comme étant formée de plusieurs dimensions (facilité d'usage, temps de réponse, etc.). Les relations entre chacune de ces dimensions (construits) et leurs items respectifs dans un questionnaire sont modélisées comme étant réflexifs. Par contre, les relations entre le concept de 2eme ordre « *qualité du système* » et ses dimensions (facilité d'usage, temps de réponse, etc.) sont modelées de manière formative. Il est donc possible de conceptualiser des construits de 2eme ordre comme des modèles réflexif-réflexif, réflexif-formatif, formatif-réflexif, et formatif-formatif. Le choix doit être justifié à l'aide d'argumentation et en se basant sur la théorie.

Dans le cas de l'exemple dans la figure 4.2, on peut facilement justifier l'existence de dimensions formatives. Un usager peut avoir accès à un système très facile à utiliser et

en même temps éprouver un temps de réponse très long. Facilité d'utilisation et le temps de réponse ne mesure pas des concepts qui sont fortement corrélés ensemble. Étant donné que les dimensions de la qualité du système ne sont pas théoriquement corrélés, cela justifie une conceptualisation formative entre le construit « *qualité du système* » et ses dimensions.

Dans la figure 4.2, les items qui mesurent les construits de 1^{er} ordre (par ex., la facilité d'usage) sont reflectifs. Ceci est très courant dans les recherches en SI/TI. Lors de nos analyses, nous chercherons alors à démontrer que ces items sont fortement corrélés ensemble. Ceci est typiquement accompli par l'analyse du Alpha de Cronbach (Pedhazur et Schmelkin, 1991) qui mesure la consistance entre des items qui, théoriquement, devraient mesurer le même construit. En d'autres mots, l'alpha de Cronbach nous aide à établir si les réponses des participants aux items (du questionnaire) élaboré par le chercheur afin de mesurer un certain concept sont consistantes les unes avec les autres.

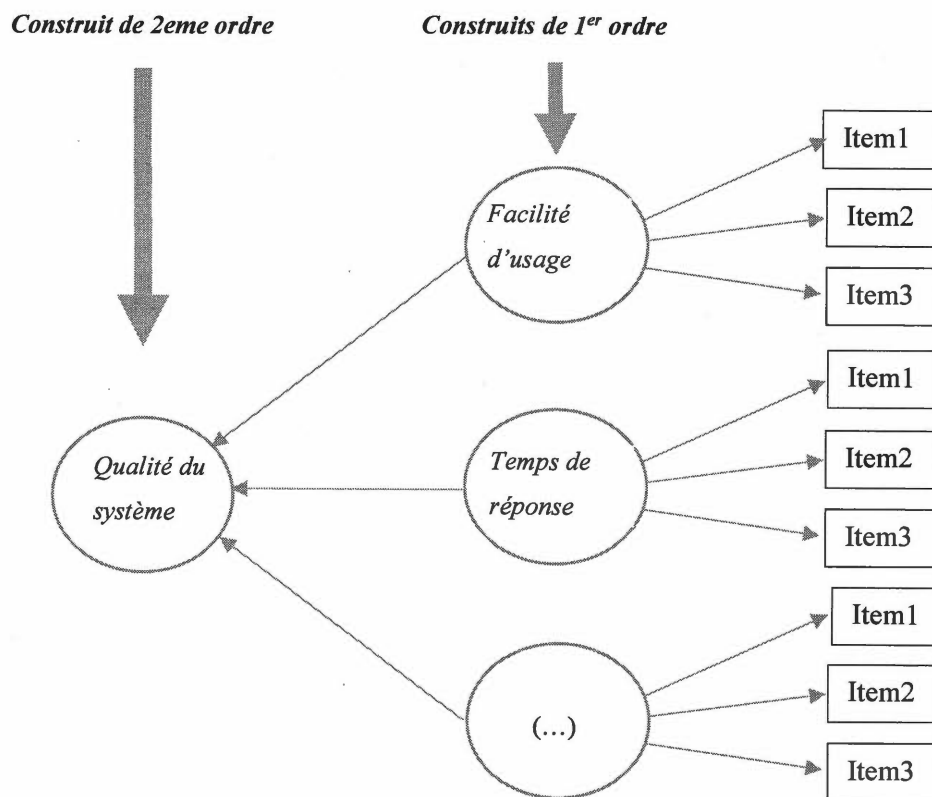


Figure 4.2 – Modélisation combinant des aspects formatifs et réflexifs

4.4.4 Description de notre modèle sous PLS-SEM

Nous avons défini notre modèle conceptuel (voir figure 3.1) en modèle structurel dans SmartPLS en utilisant un mélange de construits formatifs et réflectifs. Voir figure 4.3 ci-dessous. La partie externe du modèle structurel (démontrant les « *loadings* » entre construits et leurs items ont été définis de manière réflexive. La partie du modèle structurel décrivant les relations entre les construits est appelée « *modèle interne* ». Les relations stipulées entre ces variables latentes (construits) sont appelées « *paths* ». Cette

partie du modèle a été stipulé de manière formative. Le modèle dans la figure 4.3 dénote la « *qualité du chatbot* » comme un construit formatif de 3^e ordre. Afin de tester la validité convergente, nous avons aussi ajouté le construit « *qualité globale* » (cercle noir). Les résultats de nos analyses pour la validité convergente de la qualité du chatbot sont présentés dans la section 4.7.

Dans ce modèle, tous les construits de 2^e ordre (les dimensions de la qualité du chatbot, soient) ainsi que le construit de 3^e ordre (qualité du chatbot) ont été spécifiés en utilisant la méthode connue sous le nom de « *Repeated Indicator Approach* » (ou « *l'approche des indicateurs répétés* » – traduction libre).

Le « *Repeated Indicator Approach* » est la méthode la plus souvent utilisée dans les modèles testés sous PLS-SEM, principalement à cause de sa simplicité (Wilson & Henseler, 2007; Wetzels, Odekerken-Schroder et Van Oppen, 2009). Les indicateurs des construits de 1^{er} ordre sont réutilisés pour la construction des 2^e et 3^e ordres. Dans cette approche, l'ensemble des items utilisés pour les construits communs du 1^e ordre sont répété au 2^e. Par exemple, lorsqu'on va spécifier le construit de 2^e ordre « *Qualité de l'interaction* », il faut lui ajouter tous les items de « *Convivialité* » et de « *Personnalisation* » (comme des items réfléchifs). En d'autres termes, « *Qualité de l'interaction* » héritera de tous les sept (7) items de ses deux dimensions. De la même manière, pour le terme du 3^e ordre (Qualité du chatbot), celui-ci héritera de tous les items de ses dimensions. En d'autres termes, « *Qualité du chatbot* » aura 45 items réfléchifs qui y seront associés dans le modèle. Il faut noter que ces items « *hérités* » ou « *réutilisés* » ne sont pas visibles dans la figure 4.3 pour ne pas compliqué inutilement le modèle pour le lecteur. Pour aider le lecteur à visualiser ce processus, nous le (la) référons à la figure 4.4 qui décrit visuellement *l'approche des indicateurs répétés* pour un construit de 2^e ordre où H représente la construit de 2^e ordre et H₁, H₂, et H₃ sont les dimensions (construits de 1^e ordre).

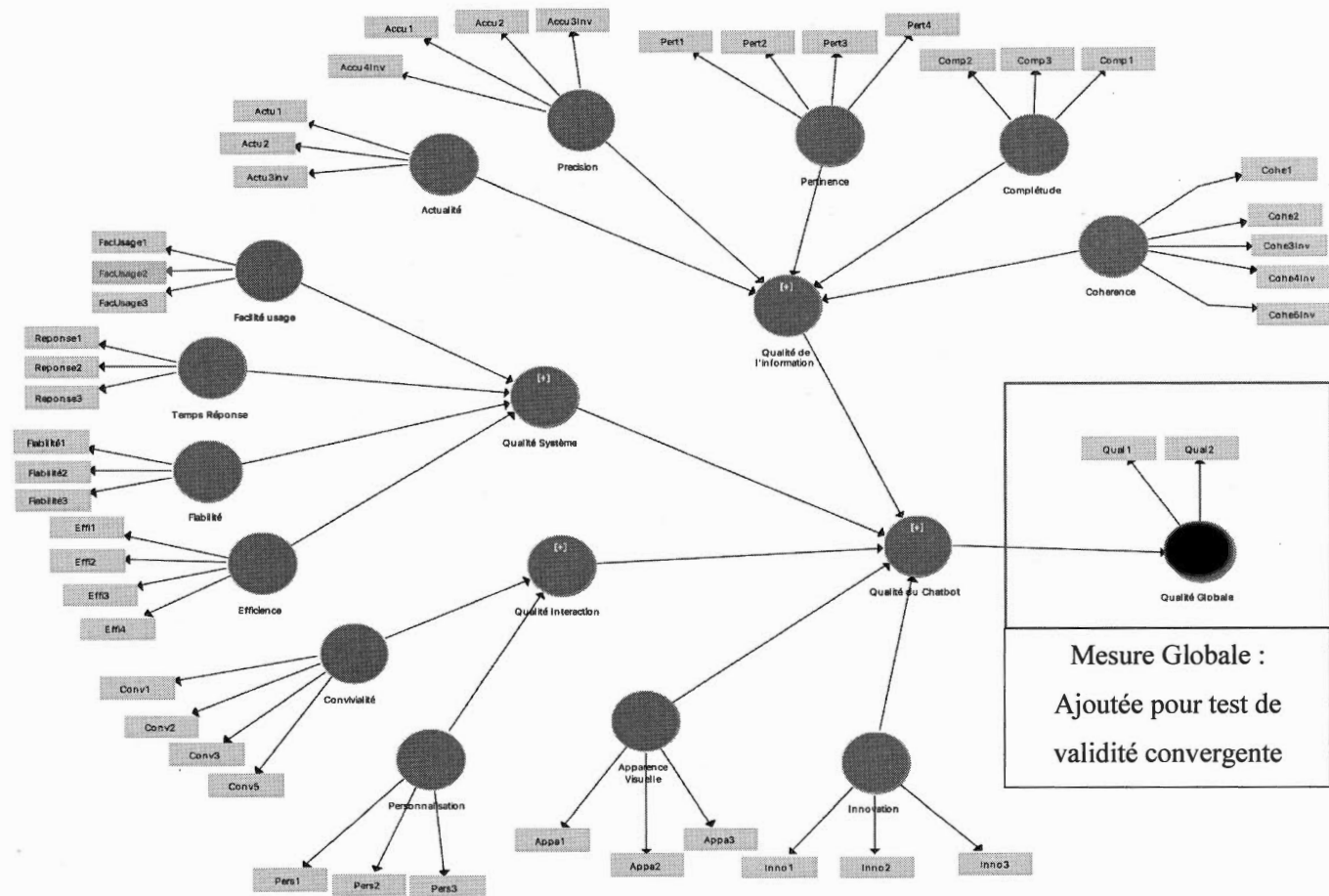


Figure 4.3 – Specification du modèle structurel initial sous SmartPLS

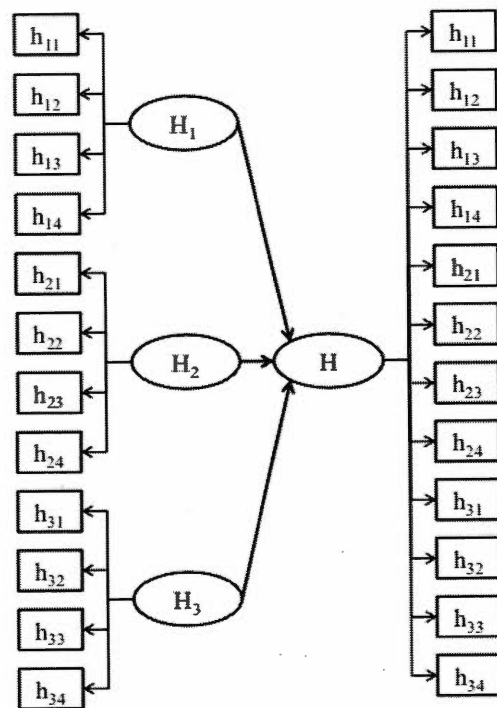


Figure 4.4 – Modélisation de construit de 2^e ordre à l’aide du « *Repeated Indicators Approach* » (Source : van Riel et coll., 2017)

4.4.5 Grandeur de notre échantillon

Le plus grand nombre de flèches arrivant vers une variable latente (construit) du modèle structurel est cinq (5) (voir Qualité du Chatbot ou Qualité de l’information). En utilisant les méthodes d’estimation de grandeur d’échantillon décrites dans la section 3.5.3, nous avons estimé qu’on aurait besoin entre 50 et 92 répondants. Notre collecte de données s’est faite en Hiver 2018 dans les classes du cours MET-1330 qui comptent chacune environ 40 à 45 étudiants. Nous avons pu obtenir $n=87$ questionnaires utilisables qui atteignait presque le montant maximum estimé ci-dessus.

4.5 Analyse du modèle externe

Fiabilité

La fiabilité représente la consistance ou la cohérence entre des mesures différentes conçues pour mesurer le même concept. La façon la plus utilisée en recherche quantitative pour établir la fiabilité des mesures est l'alpha de Cronbach. Cet indicateur de fiabilité développé par Lee Cronbach (1951) fournit une mesure de la cohérence interne d'une échelle / mesure. L'alpha de Cronbach est exprimé par un nombre entre 0 et 1. Pour établir la fiabilité des échelles dans un questionnaire, tous les alpha de Cronbach sont calculés et doivent être supérieurs ou égaux ($\geq$) à 0,7, ce qui indique une bonne fiabilité. Les résultats sont présentés dans le tableau 4.3.

En somme, les items dans le questionnaire mesurant une même variable latente sont analysés ensemble afin de générer une valeur d'alpha de Cronbach.

Validité convergente

L'objectif est de déterminer si les items utilisés pour mesurer un facteur latent commun sont réellement liés. Pour cela, nous examinons trois choses : les *Outer Loadings*, la *Composite Reliability* (CR), et l'*Average Variance Extracted* (AVE). Nous retirons les items avec des *Outer Loadings* inférieurs à 0,4. Ceux égaux ou supérieurs à 0,7 sont conservés. Les *Outer Loadings* entre 0,4 et 0,7 sont signalés et devraient être examinés avant de décider de les éliminer ou non. En les éliminant, cela augmente la *Composite Reliability* (CR) et / ou l'AVE, et les aide à les mettre dans des gammes acceptables (si elles ne le sont pas déjà) ; alors, nous pouvons les rejeter, sinon les conserver (Henseler et coll., 2009). Pour la CR, on recherche un score de 0,7 et plus. Ceci montre que les

variables convergent (Hair et coll., 2010). Les valeurs souhaitables pour l'AVE sont égales ou supérieures à 0,5 (Fornell et Larker, 1981).

Validité discriminante

La validité discriminante est démontrée lorsque chaque item de mesure est faiblement corrélé avec tous les autres construits, excepté celui auquel l'item est théoriquement associé (Gefen et Straub, 2005). Dans SmartPLS, trois indicateurs sont fournis pour aider à évaluer la validité discriminante. Ce sont respectivement des *Cross Loadings*, le critère de Fornell-Larcker, et le *Heterotrait-Monotrait Ratio* (HTMT). Cependant, une comparaison récente effectuée par Henseler, Ringle et Sarstedt (2015) de ces tests dans PLS suggère que le critère de Fornell-Larcker et l'évaluation du *Cross Loadings* peuvent ne pas toujours avoir la sensibilité adéquate pour détecter les problèmes de validité discriminante. Au lieu de cela, les auteurs recommandent d'utiliser le *Heterotrait-Monotrait Ratio*. Avec HTMT, le seuil maximum de 0,85 (Clark et Watson, 1995 ; Kline, 2011) ou de 0,90 (Teo et coll., 2008) ne devrait pas être dépassé. En somme, la validité discriminante est démontrée si le score HTMT est inférieur à 0,85/0,90.

Nous avons commencé par générer l'algorithme PLS et inspecter les *loadings* des items. Deux items (Qual3Inv et Conv4) ont été supprimés étant donné qu'ils étaient inférieurs au seuil de 0,4. L'alpha de Cronbach, la *Composite Reliability* (CR) et l'*Average Variance Extracted* (AVE) ont ensuite été évalués. Toutes les valeurs se situaient dans des limites acceptables (voir le tableau 4.3). La fiabilité et la validité convergente (pour le modèle externe) ont donc été démontrées.

Tableau 4.3 – Les résultats de SmartPLS utilisés pour déterminer la fiabilité et la validité convergente

	Alpha de Cronbach	Composite Reliability (CR)	Average Variance Extracted (AVE)
Complétude	0,813	0,889	0,728
Pertinence	0,863	0,907	0,709
Précision	0,782	0,859	0,604
Actualité	0,747	0,860	0,685
Facilité d'usage	0,831	0,899	0,749
Temps de réponse	0,940	0,962	0,893
Fiabilité	0,840	0,904	0,758
Efficience	0,838	0,892	0,677
Cohérence	0,806	0,865	0,563
Convivialité	0,813	0,878	0,645
Innovation	0,854	0,911	0,775
Personnalisation	0,810	0,887	0,723
Apparence visuelle	0,880	0,926	0,806
Qualité globale	0,866	0,962	0,893

Montrant le modèle après avoir retiré les items Conv4 et Qual3Inv, la figure 4.5 ci-dessous a été extraite à partir du logiciel SmartPLS :

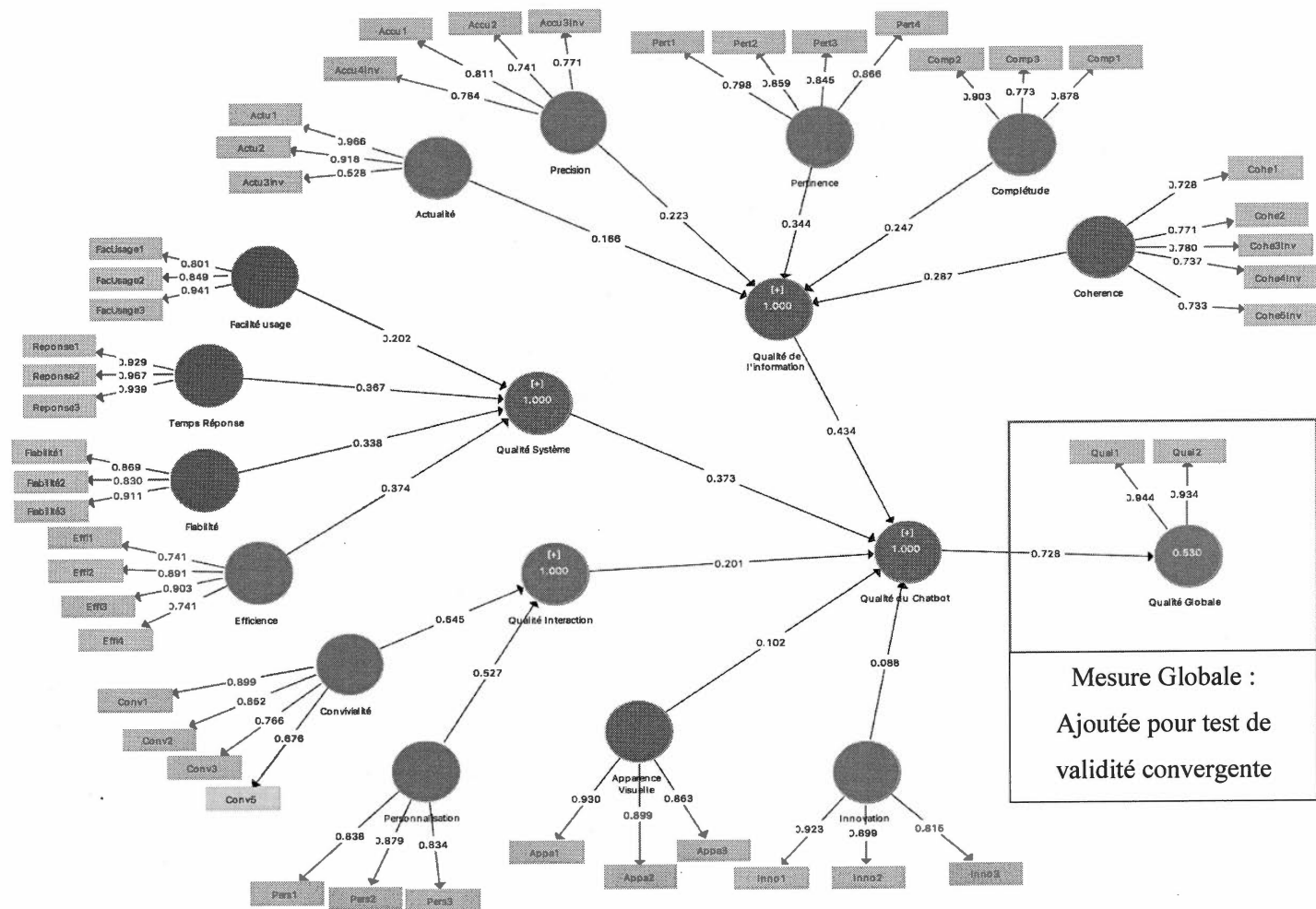


Figure 4.5 – Modèle après avoir retirés deux items (Conv4 et Qual3Inv)

Enfin, les valeurs HTMT ont été générées à l'aide de SmartPLS. Toutes les valeurs étaient inférieures à 0,85, hormis la valeur comprise (0,889) entre les dimensions « *Complétude* » et « *Pertinence* ». Nous avons décidé de combiner les éléments de ces deux dimensions dans le modèle et de les spécifier comme reflétant une variable latente que nous avons appelée « *Complétude / Pertinence* ». Nous avons encore une fois mené l'analyse PLS-SEM. Les résultats ont montré que cette action corrigeait le problème de validité discriminante. Une nouvelle matrice HTMT a été générée, et les résultats ont montré que toutes les valeurs étaient désormais inférieures à 0,85, la plus élevée étant de 0,779 entre « *Innovation* » et « *Communication personnalisée* ». Le nouveau modèle avec la combinaison des dimensions de « *Complétude* » et de « *Pertinence* » ensemble est présenté dans la figure 4.6 (ci-dessous). L'alpha de Cronbach de la nouvelle dimension « *Complétude / Pertinence* » est de 0.901 démontrant une excellente fiabilité ($>0,70$). Le *composite reliability* (CR) et le *Average Variance Extracted* (AVE) pour ce nouveau construit est de 0,922 et 0,630 respectivement. Ces deux valeurs attestent qu'il y a une excellent validité convergente.

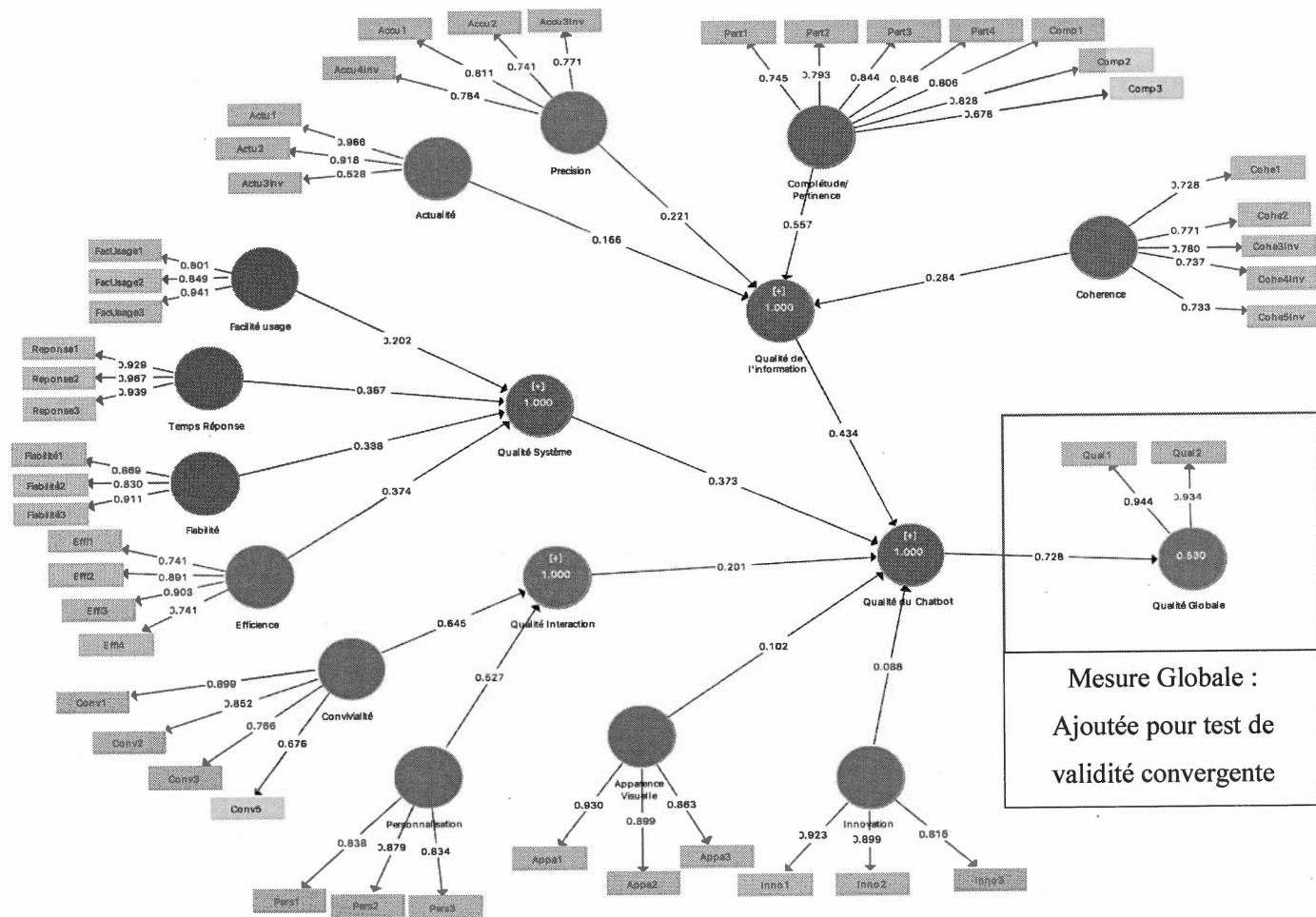


Figure 4.6 – Nouveau modèle après avoir combiné les dimensions « *Complétude* » et « *Pertinence* »

4.6 Analyse du modèle interne

Après avoir nettoyé le modèle externe, nous avons évalué le modèle interne en analysant les *Path Coefficients* (et significations statistiques) entre les dimensions de qualité théorisées et la qualité globale. L'algorithme SmartPLS a généré des *Path Coefficients*. Montrées dans le tableau suivant, les valeurs varient de +0,088 à +0,557. Toutes les valeurs sont positives. Elles sont triées de la plus grande à la plus petite (de haut vers le bas).

Dans SmartPLS, pour évaluer si les *Path Coefficients* sont significatifs ou non, des méthodes de ré-échantillonnage (comme « *Bootstrap* ») sont nécessaires. Nous avons donc utilisé Bootstrap avec les paramètres suivants : 500 sous-échantillons avec un niveau de signification de 0,05 et un type de test t bilatéral.

Le *Bootstrapping* calcule les valeurs-t pour les *Path Coefficients* (modèle interne). Ces valeurs t évaluent la signification du chemin et peuvent être utilisées pour réviser le modèle global. Nous avons déterminé la signification du *Path Coefficient* en utilisant les critères suivants : chemin significatif si $t > 1,96$ au niveau $p < 0,05$ pour les tests bilatéraux (*two-tailed*). Les résultats du « *bootstrapping* » sont présentés dans la figure 4.7 ci-dessous.

Tableau 4.4 – Valeurs des *Path Coefficients* générés par SmartPLS

De	À	<i>Path Coefficient</i> (valeur test-t)
<i>Dimensions la Qualité du Chatbot</i>		
Qualité de l'information	Qualité du Chatbot	0,434 (9,727)
Qualité du Système	Qualité du Chatbot	0,373 (9,878)
Qualité d'interaction	Qualité du Chatbot	0,201 (9,273)
Apparence visuelle	Qualité du Chatbot	0,102 (4,698)
Innovation	Qualité du Chatbot	0,088 (4,003)
<i>Sous-dimensions</i>		
Complétude / Pertinence	Qualité de l'information	0,557 (14,097)
Cohérence	Qualité de l'information	0,248 (8,333)
Précision	Qualité de l'information	0,221 (6,373)
Actualité	Qualité de l'information	0,166 (4,697)
Efficience	Qualité du Système	0,374 (10,924)
Temps de réponse	Qualité du Système	0,367 (11,772)
Fiabilité	Qualité du Système	0,338 (12,265)
Facilité d'usage	Qualité du Système	0,202 (4,387)
Convivialité	Qualité d'interaction	0,645 (10,800)
Personnalisation	Qualité d'interaction	0,527 (9,757)

Les résultats présentés dans le tableau 4.4 soutiennent toutes nos hypothèses (voir tableau 4.5 pour le récapitulatif). La qualité de l'information, la qualité de du système, la qualité de l'interaction, l'apparence visuelle, et innovation sont toutes reliées positivement et significativement à la qualité des chatbots et sont donc considérées comme des dimensions de celle-ci. Parmi ces cinq (5) dimensions, la qualité de l'information est celle avec la relation positive la plus forte avec la qualité du chatbot (path = 0,434, valeur-t = 9,727). Ceci montre que la *qualité de l'information* délivrée par un chatbot est évaluée comme étant la dimension la plus importante pour les

participants à notre étude à qui on a demandé d'évaluer la qualité du chatbot. En deuxième place, on retrouve la *qualité du système* (path = 0,373, valeur-t = 9,878). Ceci est ensuite suivi par la qualité de l'interaction entre l'utilisateur et le chatbot (path = 0,201, valeur-t = 9,273), l'apparence visuelle (path = 0,102, valeur-t = 4,698), et la perception que le chatbot est innovant (path = 0,088, valeur-t = 4,003).

Tableau 4.5 – Tableau récapitulatif – hypothèses et résultats

Hypothèse	Supportée
H1 - La qualité de l'information produite par le chatbot est positivement et significativement reliée à la qualité perçue du chatbot.	Oui
H2 - La qualité de système (du chatbot) est positivement et significativement reliée à la qualité perçue du chatbot.	Oui
H3 - La qualité des interactions entre l'utilisateur et le chatbot est positivement et significativement reliée à la qualité perçue du chatbot.	Oui
H4 - L'apparence visuelle du chatbot est positivement et significativement reliée à la qualité perçue du chatbot.	Oui
H5 - L'innovation perçue du chatbot est positivement et significativement reliée à la qualité perçue du chatbot.	Oui

Le tableau 4.4 montre aussi que les *paths* (et valeurs-t) entre les sous-dimensions de la qualité de l'information (complétude / pertinence, cohérence, précision, actualité) et la qualité de l'information sont toutes positives et significatives. La complétude / pertinence est de loin la sous-dimension de la qualité de l'information la plus forte parmi les quatre (4) sous-dimensions (path = 0,557, valeur-t = 14,097). Ceci est suivi par la cohérence (path = 0,248, valeur-t = 8,333), la précision (path = 0,221, valeur-t = 6,373) et l'actualité de l'information (path = 0,166, valeur-t = 4,697).

Pour la dimension qualité du système, la sous-dimension la plus importante est l'efficacité (path = 0,374, valeur-t = 10,924) suivie de près par le temps de réponse (path = 0,367, valeur-t = 11,772) et la fiabilité (path = 0,338, valeur-t = 12,265). Nos résultats montrent que bien que la facilité d'usage est une sous-dimension de la qualité du système (path = 0,202, valeur-t = 4,387).

Pour la qualité de l'interaction entre l'utilisateur et le chatbot, nos résultats soutiennent que la convivialité (path = 0,645, valeur-t = 10,800) et la personnalisation (path = 0,527, valeur-t = 9,757) sont toutes deux (2) reliées positivement et significativement à la qualité de l'interaction.



Figure 4.7 – Résultat du « *Bootstrapping* » montrant la signification statistique des « *paths* » du modèle

4.7 Validité convergente du construit « *qualité du chatbot* »

La validité convergente au niveau des construits peut être évaluée en examinant l'association entre une mesure du construit qui est développé et une mesure différente de ce même construit (par ex., Abernethy & Brownell, 1999). Les figures 4.6 et 4.7 démontre une relation positive (corrélation = 0,728) et significative (valeur-t = 10.250) entre la Qualité du Chatbot et la mesure de Qualité Globale du Chatbot. Cette valeur représente la corrélation entre ces construits qui est positive et significative (valeur-t > 1.96) à un niveau de $p < 0,05$ démontrant une relation relativement forte entre ces deux construits et donc supportant la validité convergente de notre mesure multidimensionnelle de qualité du chatbot.

Un deuxième indicateur qui peut être utilisé pour évaluer si nos deux mesures de qualité du chatbot sont étroitement reliées est le R^2 (ou R carré). Cette valeur représente le pourcentage (%) de la variance expliquée dans la variable dépendante par la (ou les) variable(s) indépendante(s) du modèle. Dans notre modèle, la valeur du R^2 pour la Qualité Globale (la variable dépendante) est de 0,530. Étant donné qu'il n'y a qu'une seule variable indépendante, soit la Qualité (multidimensionnelle) du Chatbot, ceci montre que notre mesure multidimensionnelle explique 53% de la variance dans la mesure globale de qualité du chatbot. D'après Wetzels et coll. (2009), l'ampleur du R^2 est petite si elle est d'environ 0,1, moyenne si elle est près de 0,25 et grande si elle est de 0,36 ou plus (voir aussi Cohen, 1988). Étant donné que la valeur du R^2 (0,530) obtenue dans notre modèle dépasse la valeur seuil de 0,36, cela nous permet donc de soutenir que notre construit de Qualité (multidimensionnelle) du Chatbot a un bon pouvoir de prévision.

4.8 Conclusion au chapitre

À partir de la revue de la littérature que nous avons menée sur les anciennes mesures de qualité de différents domaines technologiques, nous avons pu identifier les dimensions suivantes :

- La complétude des informations et la pertinence des informations adaptées du modèle de réussite du commerce électronique de DeLone et McLean (2003).
- La précision adoptée des six dimensions primaires pour l'évaluation de la qualité de données proposées par Askham et coll. (2013).
- L'actualité des informations fournis qui vient de la recherche de l'Institut canadien ICIS (2009) sur l'actualité des informations.
- La cohérence des communications provenant de la recherche d'Agar et Hobbs (1982) sur la cohérence et l'analyse des entretiens ethnographiques, ainsi que de la recherche de Louwerse et Gasser (2005) sur la cohérence et l'interprétation du discours.
- La facilité d'usage, le temps de réponse, la confiance / vie privée, l'innovation, et la personnalisation / communication sur mesure adoptées du modèle de qualité des sites Web WEBQUAL proposé par Loiacono et coll. (2002).
- La fiabilité et la convivialité qui viennent de la recherche de Shahin (2006) sur la qualité de service.
- L'efficience issue de recherches telle que celle d'Abd El-Aleem (2007) sur les sites Web.
- L'apparence visuelle qui vient de la recherche de Kuligowska (2015) sur l'évaluation de la performance de chatbots.
- L'innovation perçue du chatbot qui a été adaptée de travaux sur la qualité du site Web (Loiacono et coll., 2002) et qui est définie comme étant la créativité et l'unicité du chatbot.

CHAPITRE V

SOMMAIRE DES RESULTATS, IMPLICATIONS ET CONCLUSION

5.1 Sommaire des résultats

Notre revue de la littérature démontre que la qualité occupe une place importante en tant que mesure dans le domaine des affaires (Brady et Cronin, 2001 ; Parasuraman et coll., 1985, 1988, 1991) et celui des technologies de l'information (DeLone et McLean, 2003 ; Elling et coll., 2012 ; Loiacono et coll., 2002, 2007). Au fil du temps, les chercheurs ont développé des standards, des normes et des modèles de qualité pour évaluer différents aspects technologiques (qualité des systèmes de commerce électronique, qualité des sites Web, qualité des logiciels, etc.). Cette même revue de la littérature révèle que peu de recherches ont été menées sur l'évaluation de la qualité dans le contexte de l'utilisation des chatbots pour des fins commerciales. Cette étude a cherché à combler ce manque en développant et en validant un instrument de mesure pour la qualité perçue des chatbots. Notre étude cherchait aussi à savoir quelles sont les dimensions de qualité qui sont les plus importantes pour les usagers de ces technologies. À notre connaissance, ceci représente la première initiative dans ce genre.

Une collecte de données a ensuite été menée auprès d'étudiants au baccalauréat en administration de l'Université du Québec à Montréal, et des analyses statistiques ont été réalisées à l'aide du logiciel SmartPLS 3.0. Cela avait pour but de valider, de manière empirique, les dimensions qui influaient sur la perception de la qualité des utilisateurs de chatbots, ainsi que de les classer par ordre d'importance. Nous avons ainsi pu développer une mesure multidimensionnelle de la qualité des chatbots qui peut être utilisée comme outil pour aider les développeurs à évaluer la qualité des chatbots. Ce modèle peut être utilisé de deux manières différentes : (1) en ciblant la qualité lors

du développement d'un nouveau chatbot, (2) en diagnostiquant un chatbot déjà existant dans le but d'améliorer sa qualité.

Selon notre revue de la littérature sur le développement de nouvelles mesures, la fiabilité peut être vérifiée à l'aide du coefficient alpha de Cronbach, qui doit être supérieur à 0,70 pour prouver sa fiabilité (Pedhazur et coll., 1991). Les coefficients alpha de notre modèle respectent ce seuil, démontrant donc la fiabilité du modèle.

Afin d'identifier les dimensions de la qualité des chatbots telles que perçues par les utilisateurs de ces systèmes, nous avons utilisé un processus itératif. Notre modèle conceptuel initial (qui contenaient des dimensions extraites de la littérature sur la qualité des services, la qualité des sites Web, et d'autres sources scientifiques reliées à la qualité de la délivrance des services aux consommateurs) a été évalué en utilisant la méthode statistique appelée *Partial Least Square – Structured Equation Modeling*, (souvent appelée PLS-SEM, tout simplement).

5.2 Implications théoriques

Cette étude étend la recherche sur la qualité en développant et en validant un modèle de qualité supérieure pour les chatbots selon cinq dimensions principales (qualité de l'information, qualité du système, qualité de l'interaction, apparence visuelle et innovation). Pour la qualité de l'information, quatre sous-dimensions ont été identifiées, précisément, la cohérence, la complétude / pertinence, la précision et l'actualité des informations dans cet ordre d'importance. Pour la qualité du système, les sous-dimensions identifiées dans notre étude sont l'efficacité, le temps de réponse, la fiabilité, et la facilité d'usage (dans cet ordre d'importance). Pour la qualité de l'interaction entre l'utilisateur et le chatbot, notre étude révèle l'importance des sous-dimensions convivialité et personnalisation (dans cet ordre). Finalement, deux dimensions de plus sont ressorties; la l'apparence visuelle et l'innovation (perception

par l'utilisateur que le chatbot est innovant). Notre mesure contribue à la recherche de plusieurs manières. Premièrement, elle définit ces construits (dimensions et sous-dimensions) et deuxièmement, elle identifie 45 items de questionnaire qui servent à prédire la qualité d'un chatbot. Notre mesure pourra servir à d'autres chercheurs intéressés à faire des recherches sur la qualité des chatbots et mieux comprendre la qualité perçue du chatbot d'une compagnie peut affecter d'autres concepts importants identifiés dans la littérature tels que la satisfaction, la confiance et l'intention de faire affaire avec la compagnie à l'avenir. Donnée la nature multidimensionnelle de notre modèle, celui-ci offre aux chercheurs une opportunité d'explorer comment, non seulement la qualité des chatbots (globale) affecte d'autres concepts mais aussi comment les dimensions et sous-dimensions individuelles de la qualité des chatbots interagissent avec des concepts tels que la satisfaction et plusieurs intentions comportementales (intention d'usage, intention d'achat, etc.).

5.3 Implications pratiques

Les résultats obtenus dans cette recherche permettront aux gestionnaires de mieux comprendre comment les clients évaluent la qualité des chatbots. Les résultats suggèrent que les compagnies employant cette technologie et les concepteurs de chatbots devraient s'attarder à améliorer la qualité sur cinq (5) dimensions principales, soient : (1) la qualité de l'information fournie par le chatbot, (2) la qualité du système, (3) la qualité des interactions, (4) l'apparence visuelle (l'esthétique), et (5) l'aspect innovant du chatbot. Notre étude aide aussi à identifier des sous-dimensions de la qualité de l'information, du système ainsi que de l'interaction qui sont importants pour les usagers de chatbots. Ceci peut aider à diagnostiquer des lacunes précises au niveau de la qualité dans des chatbots existants. Par exemple, nos résultats démontrent que les perceptions de la qualité de l'information produite par un chatbot peut être augmentée en améliorant la complétude et pertinence des réponses offertes par un chatbot, ou encore, en améliorant la cohérence des réponses offertes par le chatbot. De même, la

perception qu'ont les usagers de la qualité du système pourrait être améliorée en mettant plus d'emphasis sur l'efficience du système ainsi que sa facilité d'usage. En somme, notre modèle offre un cadre qui peut servir à l'analyse et au diagnostic de chatbots existant et même à la conception de nouveaux chatbots en identifiant des aspects différents et importants de la qualité des chatbots.

5.4 Implications managériales

Les résultats de notre étude peuvent être utiles aux entreprises qui cherchent à améliorer la gestion du service à la clientèle à l'aide de nouvelles technologies. Les chatbots deviennent, de plus en plus, une technologie indispensable pour les entreprises qui désirent améliorer l'expérience client (disponibilité 24 hrs sur 24, 7 jours par semaine) tout en diminuant les coûts de la main d'œuvre. Ces chatbots doivent être bien conçus pour qu'ils puissent, non seulement servir un grand nombre de clients (quantité), mais aussi à leur offrir un service excellent (qualité). Les résultats de notre étude peuvent ainsi aider les responsables administratifs du secteur des services à la clientèle, à identifier quels sont les dimensions de la qualité des chatbots que leur clientèle apprécie le plus (ou le moins). En d'autres mots, notre mesure de qualité multidimensionnelle peut servir comme outil de diagnostic pour des chatbots existants.

5.5 Limites

Dans cette étude, il existe plusieurs limites. Premièrement, lors de la collecte des données, nous avons demandé aux répondants d'interagir avec un (1) parmi quatre (4) chatbots. Il serait important de reproduire nos résultats avec un plus grand nombre de chatbots afin d'augmenter la généralisabilité de nos résultats. Deuxièmement, notre étude a employé une méthodologie de recherche transversale (« *cross-sectional* »). Bien que cette approche méthodologique soit couramment utilisée en recherche, elle a

des limitations. Les études transversales ont certains avantages (généralement rapides, faciles à faire et relativement peu coûteuses). Elles sont très souvent basées sur des enquêtes qui utilisent des questionnaires. Les études transversales peuvent donc être sujet à certains biais. Par exemple, il se peut que les participants qui consentent à participer à notre étude sur les chatbots (étudiants au baccalauréat dans un cours en administration) diffèrent de ceux qui utiliseraient cette technologie dans un contexte réel de tous les jours. En somme, les collectes de données dans les salles de cours risquent toujours de donner un échantillon non représentatif de la population.

Une autre limite de la recherche quantitative de nature transversale est que ce type de recherche se concentre sur les relations entre les variables sans prendre compte que ces relations peuvent fluctuer ou changer à travers le temps. En d'autres mots, les études qui adoptent une approche transversale présument des relations statiques au lieu de dynamiques entre les variables. Dans le cas de notre étude, les résultats obtenues dans notre étude ne reconnaissent pas que l'importance des dimensions et des sous-dimensions peuvent changer à travers le temps.

5.6 Recherches futures

La méthode statistique Partial Least Square (PLS) permet de construire des modèles prédictifs. Cette méthode est particulièrement utile lorsque les variables indépendantes du modèle sont nombreuses et hautement colinéaires (Tobias, 1996 ; Kock, 2016). Cette méthode d'analyse des données semble être un choix judicieux dans des études sur l'identification et la validation de mesures multidimensionnelles contenant des dimensions qui pourraient être vraisemblablement corrélées (étant donné qu'elle agissent sur le même concept, la qualité du chatbot). Néanmoins, notre étude comporte des limites et soulèvent des questions qui pourraient être investigués dans des recherches futures.

En somme, nos analyses ont été effectuées en deux étapes. Premièrement, un processus itératif a été employé qui consiste à tester notre modèle à l'aide du logiciel SmartPLS 3.0 et à retirer, une à la fois, les items (indicateurs) du modèle externe qui ne sont pas suffisamment reliés à la variable latente qu'ils sont supposés refléter. Deuxièmement, une évaluation du modèle interne cherche à démontrer que les « *paths* » entre les variables latentes sont consistantes avec les hypothèses postulées (relations positives et significatives). Deux items du questionnaire utilisés pour mesurer la qualité globale des chatbots (et opérationnalisés en utilisant des échelles Likert à 5 points) étaient : « 1 » (Qual1 = Ce chatbot est de bonne qualité) et « 2 » (Qual2 = Ce chatbot est bien conçu) ont été utilisés pour évaluer la validité convergente de notre mesure multidimensionnelle de qualité du chatbot. Une corrélation significative de ,724 entre la mesure multidimensionnelle de qualité du chatbot et la mesure globale de qualité du chatbot montre qu'il y a validité convergente. Par contre, on espérait obtenir une corrélation plus haute (>.8 ou même >.9). Une corrélation significative de ,724, bien que respectable, suggère qu'il existe vraisemblablement d'autres variables (dimensions) qui pourraient être incluses dans notre modèle pour l'enrichir et le rendre plus utiles à des concepteurs de chatbots et comme outil de diagnostic pour des chatbots existants dans le but de les améliorer. À l'avenir, de nouvelles recherches de nature plus exploratoires pourraient aider à les identifier. Des entrevues semi-structurées avec des usagers réguliers de chatbots pourraient être utile pour arriver à cette fin.

À l'avenir, examiner notre modèle à l'aide d'une approche longitudinale viendrait adresser certaines limites associées avec l'approche transversale adoptée dans notre étude. Une étude longitudinale pourrait être utilisée pour évaluer si les perceptions des utilisateurs relativement à l'importance des dimensions peuvent changer à travers le temps. Il se pourrait, par exemple, que la dimension « *facilité d'usage* » soit particulièrement importante lors des premières interactions entre l'utilisateur et le chatbot. Une fois que l'utilisateur s'habitue à l'interface d'un chatbot, il est fort possible que

l'importance de cette dimension diminue dans l'évaluation de qualité des chatbot. Aussi, durant la collecte de données, nous avons demandé aux participants d'interagir avec un chatbot (choisi aléatoirement) par le participants pendant une période d'environ 15 minutes. Ceci pourrait expliquer pourquoi les « *paths* » entre les sous-dimensions « *précision* » (.221) et « *actualité* » (.166) et la qualité de l'information sont relativement basses comparativement à l'effet de la dimension « *complétude / pertinence* » (.557). En effet, il serait difficile pour un répondant de prendre en compte la précision et l'actualité de l'information fournies par le chatbot dans son évaluation de qualité du chatbot quand celles-ci ne peuvent pas être validées. Par exemple, si le chatbot recommande à l'utilisateur un certain restaurant étant donné des critères spécifiques (prix, type de nourriture, ambiance, etc.), il serait difficile pour cet usager d'évaluer ou de valider la précision et l'actualité de ces informations à moins de visiter le restaurant en question. Conséquemment, notre approche méthodologique pourraient contribuer à sous-estimer l'importance de certaines dimensions et sous-dimensions de notre modèle. Afin d'adresser cette lacune, on pourrait possiblement reproduire notre étude avec des usagers qui ont déjà eu plusieurs expériences avec un certain chatbot et peuvent supporter, avec des expériences vécues, l'importance de certaines dimensions telles que la précision des informations, l'actualité des informations, et la fiabilité à long-terme du chatbot.

RÉFÉRENCES

- Abernethy, M. A., & Brownell, P. (1997). Management control systems in research and development organizations: The role of accounting, behavior and personnel controls. *Accounting, Organizations and Society*, 22 (3-4), 233-248
- Abernethy, M. A., & Brownell, P. (1999). The role of budgets in organizations facing strategic change: An exploratory study. *Accounting, Organizations and Society*, 24(3), 189-204
- Abd El-Aleem A.K., Abd El-Wahed, W.F., Ismail, N.A. et Torkey, F.A. (2007). Efficiency Evaluation of E-Commerce Websites, World Academy of Science, Engineering and Technology, *International Journal of Electrical and Computer Engineering*, 1(4), 988-991
- Admoni, H. et Scassellati, B. (2012). A Multi-Category Theory of Intention. In : *Proceedings of COGSCI*, 1266-1271. Disponible sur : <https://mindmodeling.org/cogsci2012/papers/0227/paper0227.pdf>. Dernière visite : 29 juillet 2018.
- Alexandris, A., Kouthouris, C. et Meligdis, A. (2006). Increasing Customers' Loyalty in a Skiing Resort : The Contribution of Place Attachment and Service Quality, *International Journal of Contemporary Hospitality Management*, 18(5). 414-425.
- Al-Mamary, Y.H., & Shamsuddin, A., & Nor Aziati, A.H. (2014). Key factors enhancing acceptance of management information systems in Yemeni companies, *Journal of Business and Management Research*, 5, 108-111.
- Alotaibi, E., Al-Sabbahy, H. et Lockwood, A. (2011). Interaction Quality in Service Encounter : Scale Development and Validation, *International CHRIE Conference Refereed Track*. AmherstbScholarWorks@UMass Amherst. University of Massachusetts. Disponible sur : <https://scholarworks.umass.edu/cgi/viewcontent.cgi?referer=https://www.google.com.sg/&httpsredir=1&article=1748&context=refereed>. Dernière visite : 29 juillet 2018.
- Arazy, O. & Kopak, R. (2011). On the measurability of information quality, *Journal of the American Society for Information Science and Technology*, 62(1), 89-99

- Askham, N., Cook, D., Doyle, M., Fereday, H., Gibson, M., Landbeck, U., Lee, R., Maynard, C., Palmer, G. et Schwarzenbach, J. (2013). The Six Primary Dimensions for Data Quality Assessment, *Defining Data Quality Dimensions*, DAMA UK Working Group. Disponible sur : https://www.whitepapers.em360tech.com/wp-content/files_mf/1407250286DAMAUKDQDimensionsWhitePaperR37.pdf. Dernière visite : 29 juillet 2018.
- Ayah A., Shaidah J., Sufyan A. (2018). An Efficient Search for Context-Based Chatbots. *2018 8th International Conference on Computer Science and Information Technology (CSIT)*. ISBN: 978-1-5386-4151-4.
- Bagozzi, R.P. (1979). The Role of Measurement in Theory Construction and Hypothesis Testing : Toward a Holistic Model. In : *Conceptual and Theoretical Developments in Marketing*, O.C. Ferrell, S.W. Brown et C.W. Lamb, Jr. (eds.), American Marketing Association, Chicago, IL., 27(3), 459-489.
- Barnes, S.J. et Vidgen, R.T. (2002). An Integrative Approach to the Assessment of E-Commerce Quality, *Journal of Electronic Commerce Research*, 3(3), 114-127.
- Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science*, 45(3), 149-159.
- Bayerque, N. (2016). "VentureBeat" : A Short History of Chatbots and Artificial Intelligence. Disponible sur : <http://venturebeat.com/2016/08/15/a-short-history-of-chatbots-and-artificial-intelligence/>, Dernière visite : 14 mars 2017.
- Brady, K. et Cronin, J. (2001). Some New Thoughts on Conceptualizing Perceived Service Quality : A Hierarchical Approach, *Journal of Marketing*, 65(3). 34-49.
- Burger, J. et coll. (2001). *Issues, Tasks and Program Structures to Roadmap Research in Question & Answering (Q&A)*. New York, 1-35. Disponible sur : https://www.researchgate.net/profile/Rob_Gaizauskas/publication/228770683_Issues_tasks_and_program_structures_to_roadmap_research_in_question_answering_QA/links/00b7d52a0a1ffe6b990000000/Issues-tasks-and-program-structures-to-roadmap-research-in-question-answering-Q-A.pdf?origin=publication_detail. Dernière visite : 29 juillet 2018.
- Buttle, F. A. (1998). Word of mouth: understanding and managing referral marketing. *Journal of strategic marketing*. 6. 241-254.
- Callejas, Z., Griol, D. et López, C. R. (2011). Predicting User Mental States in Spoken Dialogue Systems, *EURASIP Journal on Advances in Signal Processing*, 6, 1-21.
- Clark, L.A. et Watson, D. (1995). Constructing Validity : Basic Issues in Objective Scale Development, *Psychological Assessment*, 7(3), 309-319.

- Cronbach, L. (1951). Coefficient Alpha and the Internal Tune of Tests. *Psychometrika*, 16, 297-334.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, Hillside, NJ: Lawrence Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Connolly, S. L., Miller, C. J., Koenig, C. J., Zamora, K. A., Wright, P. B., Stanley, R. L., et Pyne, J. M. (2018). Veterans' Attitudes Toward Smartphone App Use for Mental Health Care: Qualitative Study of Rurality and Age Differences, *JMIR Mhealth Uhealth*, 6(8). Disponible en ligne: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6125614/>. Dernière visite : 20 Février, 2019
- Daeninck, S. (2002). *Tutorial on Conversational Agents*, University of Calgary, Course name : Agent Based Software Engineering, Course SENG 609.22.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical model. *Management Science*, 35(8), 982-1003.
- Deemter, V.K., Krenn, B., Piwek, P., Klesen, M., Schroeder, M. et Baumann, S. (2008). Fully Generated Scripted Dialogue for Embodied Agents, *Artificial Intelligence Journal*, 172(10), 1239-1243.
- DeLone, W.H. et McLean, E.R. (2003). The DeLone and McLean Model of Information Systems Success : A Ten-Year Update, *Journal of Management Information Systems*, 19(4), 9-30.
- Duffy, B.R. (2002). Anthropomorphism and Robotics. *the Symposium on Animating Expressive Characters of Social Interactions*, Imperial College, London. Disponible sur : https://www.researchgate.net/profile/Brian_Duffy3/publication/228592343_Anthropomorphism_and_Robotics/links/0046352b0b04de1b74000000/Anthropomorphism-and-Robotics.pdf?origin=publication_detail. Dernière visite : 29 juillet 2018.
- Elkins, A.C. et Derrick, D.C. (2013). The Sound of Trust : Voice as a Measurement of Trust During Interactions with Embodied Conversational Agents, *Group Decision and Negotiation*, 22(5), 910-912.
- Elling, S., Leo, L., Menno, J. et Huub, V.D.B. (2012). Measuring the Quality of Governmental Websites in a Controlled Versus an Online Setting with the Website Evaluation Questionnaire, *Government Information Quarterly*, 29(3), 383-393.

- Erik F., Morten H., Kasper H. (2000). Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?. *CHI 2000*. 2(1). 345-352.
- Evanschitzky, H., Iyer, G.R., Hesse, J. et Ahlert, D. (2004). E-Satisfaction : A Re-Examination, *Journal of Retailing*, 80(3), 237-247.
- Far, B. H. (2003). Agent Based Software Engineering. *Tutorial Report for the Course : SENG 609.22*, University of CALGARY, Alberta, Canada.
- Federal Trade Commission (FTC) (2012). *Auto Repair Basics*. Disponible sur : <https://www.consumer.ftc.gov/articles/0211-auto-repair-basics#heading>. Dernière visite : 14 février 2018.
- Fink, J. (2012). Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. *4th International Conference (ICSR)*. 199-208.
- Fong, T., Nourbakhsh, I. et Dautenhahn, K. (2003). A Survey of Socially Interactive Robots, *Robotics and Autonomous Systems*, 42(3-4), 143-166.
- Fornell, C.G. et Larcker, D.F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error, *Journal of Marketing Research*, 18(1), 39-50.
- Friedlein, A. (2003). Maintaining and Evolving Successful Commercial Web Sites, *Elsevier*, 241-252.
- Garvin, D.A. (1987). Competing on the Eight Dimensions of Quality, *Harvard Business Review*, 65(6), 101-109.
- Gaudio P. et Kater K. (2000). A Life-WebGuide : An Intelligent User Interface for Web Site Navigation, *Proceedings of the 5th International Conference on Intelligent User Interfaces*, IUI. New York. 122-123.
- Gefen, D. et Straub, D. (2005). A Practical Guide To Factorial Validity Using PLS-Graph : Tutorial And Annotated Example. *Communications of the Association for Information Systems*. 16, 91-109.
- Glassman, B. (2017). What Does A Financial Advisor Do ?, *Forbes Magazine*, Disponible sur : <https://www.forbes.com/sites/advisor/2017/02/08/what-does-a-financial-advisor-do/#6a7ee6554994>, Dernière visite : 15 février 2018.
- Guesmia, S.E., Tomiuk, D. et Marcotte, S. (2018). Developing a Measure of Quality for Web-based Chatbots - Preliminary Results, *2018 International Conference of the Association of Global Management Studies*, June 20-21, Concordia University – Sir William Campus.

- Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. (2010). *Multivariate Data Analysis*. 7th Edition, Prentice Hall, Upper Saddle River, New Jersey.
- Hair, J. F., Hult, G. T. M., Ringle, C. M. & Sarstedt, M. (2013). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 1st edition. Thousand Oaks: Sage.
- Hausman, A. (2003). Professional Service Relationships : A Multicontext Study of Factors Impacting Satisfaction, Re-Patronization, and Recommendations, *Journal of Services Marketing*, 17(3), 226-242.
- Henseler, J., Ringle, C.M. et Sarstedt, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling, *Journal of the Academy of Marketing Science*, 43(1), 115-135.
- Henseler, J., Ringle, C.M. et Sinkovics, R.R. (2009). The Use of Partial Least Squares Path Modeling in International Marketing, *Advances in International Marketing*, 20, 277-319.
- Herdman, T.H. et Kamitsuru, S. (2015-2017). *NANDA International Nursing Diagnoses : Definitions & Classification*, Oxford : Wiley Blackwell.
- Hirschman, L. et Gaizauskas, R. (2001). Natural Language Question Answering : The View from Here, *Journal of Natural Language Engineering*, 7(4), 275-300.
- Ibzan, E., Balarabe, F., et Jakada, B. (2016). Customer Satisfaction and Repurchase Intentions. *Developing Country Studies*, 6(2), 96-100.
- Isaac, H. et Volle, P. (2014). *E-commerce : de la Stratégie à la Mise en Œuvre Opérationnelle* (3^e édition), Pearson Education.
- Kamsties, E. (2001). Surfacing Ambiguity in Natural Language Requirements, *PhD thesis*. Disponible sur : <https://cs.uwaterloo.ca/~dberry/HTML.documentation/AmbiguityTalk.pdf>. Dernière visite : 30 juillet 2018.
- Kandari, J. (2010). Information Quality on the World Wide Web : A User Perspective, *Thesis of the Degree of Master of Science*, University of Nebraska (Lincoln).
- Kline, R.B. (2011). *Principles and Practice of Structural Equation Modeling*. New York : Guilford Press.
- Kock, N. (2016). Non-Normality Propagation Among Latent Variables and Indicators in PLS-SEM Simulations, *Journal of Modern Applied Statistical Methods*, 15(1), 299-315.

- Kolb, C.F et Millican, P. (2006). Connecting with Elizabeth : Using Artificial Intelligence as a Data Collection Aid, Connections, *MRS Annual Conference*, 22-24 March 2006, Barbican London.
- Kramer, J., Noronha, S. et Vergo, J. (2000). A User-Centered Design Approach to Personalization, *Communications of the ACM*, 43(8), 44-8.
- Kuligowska, K. (2015). Commercial Chatbot : Performance Evaluation, Usability Metrics and Quality Standards of Embodied Conversational Agents, *Professionals Center for Business Research*, 2(2), 1-16.
- Lalwani T., Bhalotia S., Pal A., Bisen S., et Rathod V., (2018), Implementation of a Chat Bot System using AI and NLP, *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)*, 6(3), 26-30.
- Law, K.S., Wong, C.S., et Mobley, W. H. (1998). Toward a taxonomy of multidimensional constructs, *Academy of Management Review*, 23, 741-755.
- Li, S. (2011). Beyond Question Answering : Understanding the Information Need of the User, *Submitted for the Degree of Doctor of Philosophy*, University of York. Department of Computer Science.
- Loiacono, E., Watson, R. et Goodhue, D. (2007). Webqual : An Instrument for Consumer Evaluation of Web Sites, *International Journal of Electronic Commerce*, 11(3), 51-87.
- Loiacono, E., Watson, R. et Goodhue. D. (2002). WebQual™ : A Web Site Quality Instrument, *American Marketing Association : Winter Marketing Educators' Conference*, Austin, Texas : 432-438.
- Losavio, F., Chirinos, L., Lévy, N. et Ramdane, C.A. (2003). Quality Characteristics for Software Architecture, *Journal of Object Technology*, 2(2), 133-150.
- Meuter, M.L., Ostrom, A.L., Roundtree, R.I. et Bitner, M.J. (2000). Self-Service Technologies : Understanding Customer Satisfaction with Technology-Based Service Encounters, *Journal of Marketing*, 64(3), 50-64.
- Miller, H. (1996). The Multiple Dimensions of Information Quality, *Information Systems Management*, 13(2), 79-82
- Moore, G. et Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation, *Information Systems Research*, 2(3), 192-222.

- Oberman, L.M., McCleery, J.P., Ramachandran, V.S. et Pineda, J.A. (2007). EEG Evidence for Mirror Neuron Activity During the Observation of Human and Robot Actions : Toward an Analysis of the Human Qualities of Interactive Robots, *Neurocomputing*, 70(13-15), 2194-2203.
- Pacanowsky, M. et O'Donnell-Trujillo, N. (1983). Organizational Communication as Cultural Performance. *Communication Monographs*, 50(June), 126-147.
- Parasuraman, A., Berry, L. et Zeithaml, V. (1991). Refinement and Reassessment of the SERVQUAL Scale, *Journal of Retailing*, 67(4), 420-450.
- Parasuraman, A., Zeithaml, V.A. et Berry, L.L. (1985). A Conceptual Model of Service Quality and its Implication, *Journal of Marketing*, 49(4), 41-50.
- Parasuraman, A., Zeithaml, V.A. et Berry, L.L. (1988). SERVQUAL : A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality, *Journal of Retailing*, 64(1), 12-40.
- Pavlou, P.A. et Gefen, D. (2004). Building Effective Online Marketplaces with Institution-Based Trust, *Information Systems Research*, 15(1), 37-59.
- Pedhazur, E.J. et Schmelkin S.L. (1991). Measurement, Design and Analysis : An Integrated Approach, *British Journal of Mathematical and Statistical Psychology*, 45(1), 163-163.
- Peterson, R.A., Balasubramanian, S. et Bronnenberg, B.J. (1997). Exploring the Implications of the Internet for Consumer Marketing, *Journal of the Academy of Marketing Science*, 25(4), 329-346.
- Petter, S., DeLone, W. & McLean, E. (2008). Measuring information systems success: models, dimensions, measures, and interrelationships. *European Journal of Information Systems*, 17, 236-263.
- Polit, D. et Tatano Beck, C. (2010). Generalization in Quantitative and Qualitative Research : Myths and Strategies. *International Journal of Nursing Studies*, 47, 1451-1458.
- Price, R., Neiger, D. and Shanks, G. (2008), Developing a Measurement Instrument for Subjective Aspects of Information Quality, *Communications of the Association for Information Systems*, 22 (3), 49 - 74
- Radziwill, N. et Benton, M. C. (2017). Evaluating Quality of Chatbots and Intelligent Conversational Agents, *Software Quality Professional*, 19(3), 25-36.

- Rahman A., Mamun A., Islam A., (2017), Programming challenges of Chatbot: Current and Future Prospective, *Conference: 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, 75-78.
- Reeves, B. et Nass, C. (1996). *The Media Equation : How People Treat Computers, Television, and New Media Like Real People and Places*, Cambridge University Press New York.
- Rieh, S. Y. & Belkin, N. J. (1998). Understanding judgment of information quality and cognitive authority in the WWW. *Proceedings of the ASIS Annual Meeting*, 35, 279-289.
- Sarkar, M.B., Butler, B. et Steinfield, C. (1995). Intermediaries and Cybermediaries : A Continuing Role for Mediating Players in the Electronic Marketplace, *Journal of Computer-Mediated Communication*, 1(3), 1-14.
- Shafiee, M.M. et Bazargan, N.A. (2018). Behavioral Customer Loyalty in Online Shopping : The Role of E-Service Quality and E-Recovery, *Journal of Theoretical and Applied Electronic Commerce Research*, 13(1), 26-38.
- Shahin, A. (2006). SERVQUAL and Model of Service Quality Gaps : A Framework for Determining and Prioritizing Critical Factors in Delivering Quality Services. In Partha Sarathy, V. (ed), *Service quality – An introduction*, ICFAI University, Andhra Pradesh, 117-131.
- Shen, L. (2013). Relevance and Coherence, *Journal of Language Teaching and Research*, 4(4), 852-857.
- Schamber, L. (1991). Users' criteria for evaluation in a multimedia environment. *Proceedings of the 54th ASIS Annual Meeting*, 28, 126-133.
- Siddharth, G., Deep, B., Chevelyn, D.M. et Saurabh P. (2015). An E-Commerce Website Based Chatbot, (IJCSIT), *International Journal of Computer Science and Information Technologies*, 6(2), 1483-1485.
- Sokanu. (2012). *What does a Financial Advisor do ?*. Disponible sur : <https://www.sokanu.com/careers/financial-advisor/#further-reading>, Dernière visite : 28 février 2018.
- Sotiriadou, K. (2010), Efficiency: concepts, empirical investigation and applications in Greek hospitals, *Hellenic Journal of Nursing Science*, 3(2), 45-50.
- Spreitzer, G. (1995). Psychological Empowerment in the Workplace : Dimensions, Measurement, and Validation, *Academy of Management Journal*, 38(5), 1442-1465.

- Teo, T.S.H., Srivastava, S.C. et Jiang, L. (2008). Trust and Electronic Government Success : An Empirical Study, *Journal of Management Information Systems*, 25(3), 99-132.
- Thomas, C. et Fischer, G. (1997). Using Agents to Personalize the Web, *IUI'97 Proceedings of the 2nd International Conference on Intelligent User Interfaces*, ACM, New York.
- Thompson, B. (2017). 10 Questions Your Financial Advisor Should Ask You, *Forbes Magazine*. Disponible sur : <https://www.forbes.com/sites/brianthompson1/2017/10/08/10-questions-your-financial-advisor-should-ask-you/#4960b8b8159d>. Dernière visite : 15 février 2018.
- Tobias, R.D. (1996). *An Introduction to Partial Least Squares Regression*, SAS Institute Inc., Cary, NC. Disponible sur : <https://pdfs.semanticscholar.org/da6d/977f6d16b7d491d04878a37b7b7c096dfc02.pdf>. Dernière visite : 31 août 2018.
- Tomiuk, D. (2005). The impact of site-communality on the attitudinal and behavioural components of site-loyalty : a cross-sectional study, *Thèse de Doctorat – Université McGill*. Disponible sur : http://digitool.library.mcgill.ca/R/?func=dbin-jump-full&object_id=85960&local_base=GEN01-MCG02. Dernière visite : 4 Octobre, 2018
- Unger, C. et Cimiano, P. (2011). Representing and Resolving Ambiguities in Ontology-Based Question Answering. In *Proceedings of TIWTE*, Stroudsburg, PA, USA. Association for Computational Linguistics, 15(2), 40-49.
- van Riel, A. C. R., Henseler, J., Kemény, I., Sasovova, Z. (2017). Estimating Hierarchical Constructs Using Consistent Partial Least Squares: The Case Of Second-Order Composites Of Common Factors, *Industrial Management & Data Systems*, 117(3), 459-477.
- Venkatesh, A., Khatri, C., Ram, A., Guo, F., Gabriel, R. Nagar, A., Prasad, R., Cheng, M., Hedayatnia, B., Metallinou, A., Goel, R., Yang, S. et Raju, A. (2017). On Evaluating and Comparing Conversational Agents, *31st Conference on Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA.

- Vlachos, P. A., & Aristeidis, A. (2009). Formative Versus Reflective Measurement for Multidimensional Constructs. *SSRN Electronic Journal*. Disponible sur: <https://poseidon01.ssrn.com/delivery.php?ID=133113085074066124096125069102031112006064056058069029104005031119111001112029113067098033019101022122010124088075080068107114059052014001081096121025066071123117100018049010097027098125095009103077012096080116066007099072019124101026029026093110025103&EXT=pdf>. Dernière visite : 4 Octobre, 2018.
- Vugt, V.H., Bailenson, J., Hoorn, J. et Konijn, E. (2010). Effects Official Similarity on User Responses to Embodied Agents, *ACM Transactions on Computer-Human Interaction*, 17(2), 21-22.
- Wang, R.Y. et Strong, D.M. (1996). Beyond Accuracy : What Data Quality Means to Data Consumers, *Journal of Management Information Systems*, 12(4), 5-34.
- Weizenbaum, J. (1966). ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine, *Communications of the ACM*, 9(1), 36-45.
- Westerman, D., Spence, P. R. & Van Der Heide, B. (2014). Social media as information source: Recency of updates and credibility of information. *Journal of Computer-Mediated Communication*, 19(2), 171-183.
- Wetzels, M. Odekerken-Schroder, G. Van Oppen, C. (2009). Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration. *MIS Quarterly*, 33 (1), 177-195.
- Wilbert, O.G. (2002). *An Introduction to GUI Design Principles and Techniques, The Essential Guide to User Interface Design* (Second Edition), Published by John Wiley & Sons, Inc.
- Wilson, B. and Henseler, J. (2007), "Modeling Reflective Higher-Order Constructs using Three Approaches with PLS Path Modeling: A Monte Carlo Comparison", in *Thyne, M. and Deans, K. R. (Eds.), Conference proceedings ANZMAC 2007*. ANZMAC, Dunedin, New Zealand, pp. 791-800.
- Wong, K. K. K. Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS, *Marketing Bulletin*, 24, Technical Note 1, 1-32. Disponible sur http://marketing-bulletin.massey.ac.nz/v24/mb_v24_t1_wong.pdf. Dernière visite : 7 Octobre, 2018
- Wright, R.T., Campbell, D.E., Thatcher, J.B., et Roberts, N. (2012). Operationalizing Multidimensional Constructs in Structural Equation Modeling : Recommendations for IS Research, *Communications of the Association for Information Systems*, 40(1), 367-412.

- Xu, A., Liu, Z., Guo, Y., Sinha, V. et Akkiraju, R. (2017). A New Chatbot for Customer Service on Social Media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 3506-3510. Disponible sur : https://www.researchgate.net/profile/Anbang_Xu/publication/313204805_A_New_Chatbot_for_Customer_Service_on_Social_Media/links/5a4c7220458515a6bc6c86ef/A-New-Chatbot-for-Customer-Service-on-Social-Media.pdf?origin=publication_detail. Dernière visite : 9 août 2018.
- Yun, N. (2007). *Analysis of Semantic Classes : Toward Non-Factoid Question Answering*, PhD Thesis submitted Graduate Department of Computer science, University of Toronto.
- Zeithaml, V., Parasuraman, A., and Berry, L. (1985). Problems and Strategies in Services Marketing, *Journal of Marketing*, 49, 33-46.

ANNEXE A

TABLEAU RÉCAPITULATIF DES FACTEURS DE QUALITÉ IDENTIFIÉS DANS LA REVUE DE LA LITTÉRATURE

Dimensions de qualité (Garvin, 1987)				Performance
				Caractéristiques
				Fiabilité
				Conformité
				Durabilité
				Facilité d'entretien
				Esthétique
				Qualité perçue
Concept de qualité en TI/SI ¹⁷ et Concept de qualité de service en Marketing	Qualité de l'information (DeLone et McLean, 2003)	Qualité des systèmes		Adaptabilité
				Disponibilité
				Fiabilité
				Temps de réponse
				Convivialité
		Qualité de l'information		Complétude
				Facilité de compréhension
				Personnalisation
				Pertinence
				Sécurité
		Qualité du service		Assurance
				Empathie
				Réactivité
		Utilisation		Nature de l'utilisation
				Modèles de navigation
				Nombre de visites sur le site
				Nombre de transactions exécutées
		Satisfaction de l'utilisateur		Répéter les achats
				Répéter des visites
				Enquêtes auprès des utilisateurs
		Avantages nets		Répéter les achats
				Répéter des visites
				Enquêtes auprès des utilisateurs

¹⁷ TI = Technologies de l'information. SI = Systèmes d'information

			Répéter les achats
			Répéter des visites
Qualité de système (Kuligowska, 2015)	Apparence visuelle		
	Forme de mise en œuvre sur le site Web		
	Unité de synthèse vocale		
	Base de connaissances intégrée		
	Présentation de connaissances et fonctionnalités supplémentaires		
	Capacités de conversation en contexte de la sensibilité		
	Traits de personnalité		
	Options de personnalisation		
	Réponses d'urgence dans des situations inattendues		
	Possibilité de classer le chatbot et le site Web par l'utilisateur		
Qualité de service (Parasuraman et coll., 1988)	Fiabilité		
	Tangibles		
	Réactivité		
	Assurance		
	Empathie		
Qualité de l'interaction (Alotaibi et coll., 2011)	Communication sociale		
	Compétence et assistance liées aux tâches		
Qualité des sites Web (Loiacono et coll., 2002)	Facilité d'usage	Facilité de compréhension	
	Utilité	Opérations intuitives	
		Qualité d'information	
		Ajustement de l'information à la tâche (<i>Informational Fit-to-Task</i>)	
		Interactivité	
		Confiance	

			Temps de réponse
		Divertissement	Attrait visuel
			Innovation
			Flux
		Relation complémentaire	Image cohérente
			Intégrité on-line
			Mieux que les canaux alternatifs
		Service client	Service client
Qualité et chatbots (Radziwill et Benton, 2017)	Efficacité	Performance	
	Effectivité	Fonctionnalité	
		Humanité	
	Satisfaction	Aspect affectif	
		Éthique et comportement	
		Accessibilité	

ANNEXE B

SIGNIFICATIONS DES FACTEURS DE QUALITÉ EXTRAITS DE LA LITTÉRATURE

Expressions	En anglais	Descriptions
Qualité de l'information.	<i>Information Quality</i>	Elle capture le problème du contenu du commerce électronique (DeLone et McLean, 2003).
Qualité des systèmes.	<i>System Quality</i>	Elle mesure les caractéristiques souhaitées d'un système de commerce électronique (DeLone et McLean, 2003).
Adaptabilité.	<i>Adaptability</i>	En génie logiciel, c'est la capacité du produit logiciel à être adapté à différents environnements spécifiés en utilisant seulement sa propre fonctionnalité (Losavio et coll., 2003).
Disponibilité.	<i>Availability</i>	Mesure dans laquelle l'information est physiquement accessible (Kandari, 2010).
Fiabilité.	<i>Reliability</i>	En génie logiciel, c'est la capacité du logiciel à maintenir son niveau de performance dans des conditions déclarées pendant une période déterminée (Losavio et coll., 2003).
Temps de réponse.	<i>Response Time</i>	En génie logiciel, c'est la capacité du produit logiciel à fournir un temps de réponse, un temps de traitement et des débits appropriés lors de l'exécution de sa fonction dans des conditions définies (Losavio et coll., 2003).
Complétude.	<i>Completeness</i>	Mesure dans laquelle les informations ne manquent pas et sont d'une ampleur et d'une profondeur suffisantes pour la tâche à accomplir (Kandari, 2010).
Facilité de compréhension.	<i>Ease of understanding (or Understandability)</i>	Mesure dans laquelle l'information est claire sans ambiguïté et facile à comprendre (Wang et Strong, 1996). Cité par Kandari (2010).

Personnalisation.	<i>Personalization</i>	La personnalisation est une boîte à outils de technologies et de fonctionnalités d'application utilisées dans la conception d'une expérience d'utilisateur final (Kramer et coll., 2000).
Pertinence.	<i>Relevance (or Relevancy)</i>	Mesure dans laquelle l'information est applicable et utile pour la tâche à accomplir (Wang et Strong, 1996 ; cité par Kandari, 2010).
Sécurité.	<i>Security</i>	Mesure dans laquelle l'accès à l'information est limité de façon appropriée pour maintenir sa sécurité (Kandari, 2010).
Qualité du service.	<i>Service Quality</i>	Le soutien global fourni par le fournisseur de services s'applique indépendamment du fait que cette assistance soit fournie par le département du système d'information, une nouvelle unité organisationnelle ou sous-traitée à un fournisseur de services Internet (ISP) (DeLone et McLean, 2003).
Assurance.	<i>Assurance</i>	Connaissances de courtoisie, la capacité des employés à inspirer la confiance et la confiance (Emel, 2014).
Empathie.	<i>Empathy</i>	Le soin, l'attention individualisée que l'entreprise fournit à ses clients (Emel, 2014).
Réactivité.	<i>Responsiveness</i>	Aider les clients et fournir un service rapide (Emel, 2014).
Nature de l'utilisation.	<i>Nature of Use</i>	La façon dont le client utilise le service fourni.
Modèles de navigation.	<i>Navigation Patterns</i>	Ils montrent comment construire des architectures de navigation utilisables pour les systèmes d'information Web (Gustavo et coll., 1999).
Nombre de visites sur le site.	<i>Number of site visits</i>	Le nombre de visites d'un site Web aide à identifier la valeur des informations qu'il contient.
Nombre de transactions exécutées.	<i>Number of transactions executed</i>	Le nombre de transactions exécutées aide à évaluer la demande d'un service.

Satisfaction de l'utilisateur.	<i>User satisfaction</i>	Elle demeure un moyen important de mesurer l'opinion des clients sur le système de commerce électronique et devrait couvrir tout le cycle de l'expérience client depuis la recherche d'informations jusqu'à l'achat, le paiement, la réception et le service (DeLone et McLean, 2003).
Répéter les achats.	<i>Repeat Purchases</i>	Les achats répétés permettent d'identifier la demande des clients, donc, elles permettent d'identifier la satisfaction des clients.
Répéter des visites.	<i>Repeat Visits</i>	Les visites répétées permettent d'évaluer les intentions des visiteurs au contenu.
Enquêtes auprès des utilisateurs.	<i>User Surveys</i>	Elles permettent d'améliorer les services fournis et de donner une meilleure expérience utilisateur.
Avantages nets.	<i>Net Benefits</i>	Ce sont les mesures de succès les plus importantes, car elles captent l'équilibre des impacts positifs et négatifs du commerce électronique sur les clients, les fournisseurs, les employés, les organisations, les marchés, les industries, les économies, et même les sociétés (DeLone et McLean, 2003).
Économies de coûts.	<i>Cost Savings</i>	Internet permet aux services du commerce électronique d'économiser sur les coûts de déplacement, de consultation, etc.
Marchés élargis.	<i>Expanded markets</i>	En raison du commerce électronique basé sur Internet, les marchés se sont élargis au cours du temps.
Ventes additionnelles incrémentales.	<i>Incremental Additional Sales</i>	L'incrémentation des ventes additionnelles est une conséquence de la facilité d'exploiter le commerce électronique sur Internet.
Réduction des coûts de recherche.	<i>Reduced search costs</i>	Fournir des services sur Internet évite des procédures supplémentaires non nécessaires (déplacements, consultations, etc.), cela conduit à réduire les coûts.

Économies de temps.	<i>Time Savings</i>	Internet permet aux services du commerce électronique d'économiser le temps en évitant les déplacements, les consultations, etc.
Forme de mise en œuvre sur le site Web.	<i>Form of Implementation on the Website</i>	Une entreprise qui achète un chatbot est rarement prête à reconstruire complètement son service Web existant, donc les formes d'implémentation des assistants virtuels ont tendance à être assez variées (Kuligowska, 2015).
Unité de synthèse vocale.	<i>Speech Synthesis Unit</i>	L'une des composantes de qualité d'un assistant virtuel est le module <i>Text-To-Speech</i> qui convertit le texte écrit en un discours synthétique (Deemter et coll., 2008). La recherche montre que les agents conversationnels incorporés peuvent utiliser leur voix synthétisée afin d'augmenter la confiance des utilisateurs (Elkins et coll., 2013).
Base de connaissances intégrée.	<i>Knowledge Base</i>	La base de connaissances stocke pour un assistant virtuel toutes les données, informations et connaissances sur la réalité, et par conséquent, il a une signification fondamentale pour le fonctionnement de l'assistant virtuel (Gaudio et coll., 2000).
Présentation des connaissances et fonctionnalités supplémentaires exécutées par chatbots.	<i>Presentation of Knowledge and Additional Functionalities Performed by chatbots</i>	L'évaluation de la forme exceptionnelle de présentation des connaissances, ainsi que les fonctions supplémentaires accomplies par chaque agent virtuel et les autres moyens disponibles pour faciliter la navigation des utilisateurs sur le site Web complètent l'évaluation des compétences (Kuligowska, 2015).
Capacités de conversation et contexte La sensibilité.	<i>Conversational Abilities, Language Skills and Context Sensitiveness</i>	Le plus grand défi dans la conception d'une architecture de chatbot est d'inventer le mécanisme de détection de contexte du dialogue, ce qui rend le chatbot au rythme d'un sujet de

		conversation en constante évolution (Kuligowska, 2015).
Traits de personnalité.	<i>Personality Traits</i>	Pour devenir crédible aux yeux des utilisateurs, les chatbots commerciaux doivent non seulement être dotés d'une expertise, mais aussi d'une expression de personnalité (Kuligowska, 2015).
Réponses d'urgence dans des situations inattendues.	<i>Emergency Responses in Unexpected Situations</i>	L'assistant virtuel commercial doit être capable de répondre intelligemment, diplomatiquement et patiemment aux situations d'urgence. Le chatbot devrait également reconnaître les fautes de frappe, les fautes d'orthographe ou les expressions familières utilisées dans la conversation (Kuligowska, 2015).
Possibilité de classer le chatbot et le site Web par l'utilisateur.	<i>Possibility of Rating chatbot and the Website by the User</i>	La conversation avec un assistant virtuel sert non seulement de présentation de l'offre commerciale à un client potentiel. La valeur ajoutée pour le propriétaire du chatbot est la réaction des utilisateurs sur le chatbot et / ou sur le site Web (Callejas et coll., 2011) ainsi que la volonté des clients de recommander un produit ou un service à quelqu'un d'autre (Kuligowska, 2015).
Tangibles.	<i>Tangibles</i>	Installations physiques, équipement et apparence du personnel (Shahin, 2006).
Qualité de l'interaction.	<i>Interaction Quality</i>	Elle fait référence à l'expérience d'un client résultant de l'interaction avec l'élément humain de l'organisation du service (Alexandris et coll., 2006).
Communication sociale.	<i>Sociality Communication</i>	La socialité désigne la performance qui encourage une coopérative, une douceur sociale, un vide d'interactions intenses avec les autres (Pacanowsky et coll., 1983).

Compétence et assistance liée aux tâches.	<i>Competence & Task-Related Helpfulness</i>	Il est suggéré que pour une interaction de qualité, les aspects sociaux seront insuffisants si l'employé de service ne fournit pas l'élément de service de base (Hausman, 2003). Il s'agit ici de la capacité d'un employé à exécuter une tâche avec succès (Spreitzer, 1995), et il définit dans quelle mesure le prestataire peut influencer sur le résultat de l'interaction grâce à ses compétences (DeLone et McLean, 2003).
Conséquences (satisfaction de la rencontre)	<i>Consequences (Encounter Satisfaction)</i>	Une conceptualisation d'un facteur structurel et dynamique qui affecte l'interaction sociale en général (Alotaibi et coll., 2011).
Fidélité.	<i>Loyalty</i>	La fidélité des clients envers le fournisseur et l'organisation du service, et en plus de la satisfaction du client, aide à mieux comprendre les expressions courtoises et la connexion personnelle (Alotaibi et coll., 2011).
WOM.	<i>Word of Mouth</i>	<i>Word of Mouth</i> : Le fait que les consommateurs fournissent des informations à d'autres consommateurs (Buttle, 1998).
Qualité des sites Web.	<i>Web Site Quality</i>	Elle se concentre sur la perception du consommateur de la qualité du site Web dans le contexte de la prédiction du comportement de réutilisation du site (Loiacono et coll., 2002).
Facilité d'usage.	<i>Ease of Use</i>	Le degré auquel une personne croit que l'utilisation d'un système particulier serait exempte d'efforts (Venkatesh et coll., 2003).
Opérations intuitives.	<i>Intuitive Operation</i>	Facilité d'utiliser et de naviguer (Loiacono et coll., 2002).
Utilité.	<i>Usefulness</i>	Le degré auquel une personne estime que l'utilisation d'un système particulier améliorerait son rendement au travail (Venkatesh et coll., 2003).

Ajustement de l'information à la tâche.	<i>Functional Fit-to-task</i>	Répond aux besoins de la tâche et améliore les performances (Loiacono et coll., 2002).
Interactivité.	<i>Interactivity</i>	Communication sur mesure entre les consommateurs et l'entreprise (Loiacono et coll., 2002).
Confiance.	<i>Trust</i>	Communication sécurisée et respect de la confidentialité des informations (Loiacono et coll., 2002).
Divertissement.	<i>Entertainment</i>	Elle prédit les intentions de réutiliser un site Web (Loiacono et coll., 2002).
Attrait visuel.	<i>Visual Appeal</i>	L'esthétique d'un site Web (Loiacono et coll., 2002).
Innovation.	<i>Innovativeness</i>	La créativité et l'unicité de la conception du site (Loiacono et coll., 2002).
Flux.	<i>Flow</i>	L'effet émotionnel de l'utilisation du site Web et l'intensité de l'implication (Loiacono et coll., 2002).
Relation complémentaire.	<i>Complementary Relationship</i>	Des trucs complémentaires traitant de l'ergonomie du site, des solutions alternatives, etc.
Image cohérente.	<i>Consistent Image</i>	L'image du site Web est compatible avec l'image projetée par l'entreprise à travers d'autres médias (Loiacono et coll., 2002).
Intégrité on-line.	<i>On-Line Completeness</i>	Permettre que toutes les transactions ou la plupart des transactions nécessaires soient effectuées en ligne (par exemple, un achat sur le site Web) (Loiacono et coll., 2002).
Service à la clientèle.	<i>Customer Service</i>	La réponse aux demandes de renseignements et aux commentaires des clients (Loiacono et coll., 2002).

ANNEXE C

COURRIEL DE SOLLICITATION
DES ENSEIGNANTS

Cher (chère) enseignant(e) du cours MET1330,

Dans le cadre d'un projet de mémoire pour l'obtention de la Maitrise en Informatique de Gestion à l'UQAM, j'entreprends une étude sur l'identification et validation des facteurs qui influencent les perceptions de qualité des chatbots.

J'aimerais obtenir votre consentement pour venir dans votre classe (en début de cours) pour faire du recrutement auprès de vos étudiants. **La collecte de données se fera en dehors du cours. Ma visite servira donc seulement à faire du recrutement.**

Lors de ma visite dans votre cours:

- (1) Je vous demanderai de signer une lettre de consentement me donnant votre autorisation écrite à expliquer à vos étudiants le but de ma recherche et à les inviter à y participer.
- (2) Je distribuerai à chacun de vos étudiants, une lettre d'invitation à participer à mon étude. Cette lettre contiendra (a) une courte explication de mon étude, (b) l'adresse URL du questionnaire en ligne, et (c) des détails sur un tirage auquel vos étudiants pourront s'inscrire après avoir répondu au questionnaire en ligne.
- (3) Je demanderai à vos étudiants s'ils (elles) ont des questions et je prendrai quelques minutes à répondre à leurs questions.

En tout, j'estime que **ma visite à votre cours ne prendra pas plus de 10 minutes.**

J'apprécie votre collaboration pour réaliser cette étude et je vous informe que la contribution de vos étudiants me sera très utile. Votre aide est complètement volontaire et vous et vos étudiants ne serez sujets à aucune pression à participer à mon étude. Dans l'attente d'une réponse favorable, je vous prie d'agréer l'expression de mes salutations distinguées.

Seyf Eddine Guesmia
Étudiant à la Maitrise en Informatique de Gestion
UQAM

ANNEXE D

LETTRE DE CONSENTEMENT
DES ENSEIGNANTS

Nom et Prénom : GUESMIA Seyf Eddine
Spécialité : Maîtrise en informatique de gestion (profil mémoire)
Adresse courriel : guesmia.seyf_eddine@courrier.uqam.ca
Numéro téléphone : 438-923-6320

 (date)

**Objet : Lettre de Consentement des Enseignants du MET1330 -
 Recrutement de vos Étudiants dans votre Cours**

Cher (chère) Monsieur (Madame) X,

Je sollicite votre autorisation de me présenter dans votre cours pour inviter vos étudiants à participer à mon étude. Ceci prendra environ 10 minutes de votre temps de cours. Mon étude cherche à identifier et valider les facteurs qui influencent la perception de qualité des ChatBots. Vos étudiants qui accepteront à participer à mon étude seront demandés de remplir un questionnaire en ligne en dehors des heures de votre cours. Ma visite dans votre cours servira seulement à recruter des étudiants et à répondre à vos et à leurs questions.

J'apprécie votre collaboration pour réaliser mon étude et je vous informe que la contribution de vos étudiants me sera très utile.

Seyf E. Guesmia

Je, _____ *nom de l'enseignant(e)* _____, consente à ce que M. Seyf Guesmia explique à mes étudiant le but de sa recherche et les invite à y participer. M. Seyf Guesmia m'a informé clairement de son projet de recherche. J'accorde mon consentement libre et éclairé et je comprends mon droit à retirer ce consentement en tout temps et à ne plus lui permettre de faire son recrutement lors de sa visite dans mon cours.

 (signature de

l'enseignant(e) _____

ANNEXE E

LETTRE D'INVITATION
AUX PARTICIPANTS

Cher Monsieur/Madame,

Dans le cadre d'un projet de mémoire pour l'obtention de la Maîtrise en Informatique de Gestion à l'UQAM, j'entreprends une étude sur l'identification et validation des facteurs qui influencent les perceptions de qualité des chatbots. Le but de la recherche est de développer une mesure de qualité qui aidera aux développeurs, aux entreprises, ou aux parties intéressées à développer et déployer les chatbots de qualité au niveau de leur infrastructure d'affaires. Cette mesure de qualité pourra aussi être utilisée comme outil de diagnostic afin d'évaluer et améliorer la qualité de chatbots existants. Votre participation ou non-participation à cette étude n'aura aucun effet sur votre note dans le cours MET1330. Bien que votre participation à l'enquête soit complètement volontaire, j'apprécierai énormément votre aide car elle m'est essentielle pour assurer le succès de mon projet.

Voici ce qu'il faut faire:

- 1) Allez à (*mettre le lien pour remplir le questionnaire en ligne sur « Google Forms » ici*)
- 2) Lire et accepter le consentement de participation à l'étude. Cochez les deux cases en bas de la page.
- 3) Interagir pendant 5 à 10 minutes avec un chatbot que vous allez choisir parmi la liste de chatbots.
- 4) Vous aurez ensuite à répondre à quelques questions en reflétant sur votre interaction avec le chatbot (par ex., temps de réponse, facilité d'utilisation, précision de l'information, etc.).

Tous les renseignements collectés seront totalement confidentiels et anonymes, vous ne serez pas identifié dans l'étude. Le contenu du questionnaire sera confidentiel. Seuls les résultats, qui seront sous un format agrégé suite à notre analyse, seront présentés dans mon rapport final. Aucun risque connu ou anticipé n'a été identifié. À tout moment durant le sondage, vous pouvez refuser de répondre à une question et vous retirer de l'étude, sans aucune conséquence négative. En tout, nous estimons qu'interagir avec le chatbot (5 à 10 minutes) et répondre au questionnaire (environ 10 minutes) devraient prendre de 15 à 20 minutes en tout.

À la fin du questionnaire, vous aurez l'opportunité de cliquer sur un lien qui vous dirigera à une autre page où vous aurez, si vous le voulez, la chance de vous inscrire à un **tirage (1 prix de \$30, 2 prix de \$20)**. De plus, pour les participants intéressés à recevoir le rapport de l'étude, n'hésitez pas à me contacter par courriel (voir adresse courriel ci-dessous).

GUESMIA Seyf Eddine
Candidat à la Maîtrise en Informatique de Gestion, UQAM
guesmia.seyf_eddine@courrier.uqam.ca
438-923-6320

ANNEXE F

QUESTIONNAIRE

Svp répondre aux questions suivantes en vous basant sur vos interactions avec le ChatBot. Indiquez votre niveau d'accord ou de désaccord avec les commentaires suivants:

Tout à fait en désaccord
Plutôt en désaccord
Ni d'accord ni en désaccord
Plutôt d'accord
Tout à fait d'accord

Complétude des informations :

Les réponses fournies sont complètes	1	2	3	4	5
Les informations produites par le Chatbot sont suffisantes	1	2	3	4	5
Le Chatbot répond à toutes les questions de l'utilisateur	1	2	3	4	5

Pertinence des informations:

Les informations fournies par ce Chatbot sont bonnes à savoir	1	2	3	4	5
Chaque question/commentaire/réponse était pertinente	1	2	3	4	5
Le Chatbot a su répondre à mes besoins informationnels	1	2	3	4	5
Chaque question/commentaire/réponse était utile	1	2	3	4	5

Précision:

Chaque question/recommandation provenant du Chatbot est précise	1	2	3	4	5
Chaque question/recommandation du Chatbot était claire pour moi	1	2	3	4	5
Certaines communications manquaient de précision (<i>item inversé</i>)	1	2	3	4	5
Certaines communications me semblaient ambiguës (<i>item inversé</i>)	1	2	3	4	5

Actualité des informations fournies:

L'information fournie par le Chatbot semble récente	1	2	3	4	5
Les informations produites par ce Chatbot paraissent être "à jour"	1	2	3	4	5
Certaines informations semblaient parfois désuètes (<i>item inversé</i>)	1	2	3	4	5

Cohérence des communications:

Chaque communication était cohérente avec toutes celles qui la précédaient	1	2	3	4	5
Les questions posées avaient une suite logique	1	2	3	4	5
Les commentaires du Chatbot étaient parfois contradictoires (<i>item inversé</i>)	1	2	3	4	5
Certaines communications semblaient être incohérentes (<i>item inversé</i>)	1	2	3	4	5
Les communications sautaient d'un thème à un autre (<i>item inversé</i>)	1	2	3	4	5

Facilité d'utilisation:

L'interface est intuitive	1	2	3	4	5
Ce Chatbot est facile à utiliser	1	2	3	4	5
Je me suis vite habitué à utiliser ce Chatbot	1	2	3	4	5

Temps de réponse:

Ce Chatbot est rapide	1	2	3	4	5
Le temps que ça prend au Chatbot à répondre est raisonnable	1	2	3	4	5
Le temps de réponse est acceptable	1	2	3	4	5

Aspects techniques/Fiabilité:

Ce Chatbot fonctionne bien	1	2	3	4	5
Ce Chatbot ne m'a pas causé de frustration	1	2	3	4	5
Ce Chatbot fonctionne sans problèmes	1	2	3	4	5

Efficacité:

Obtenir une réponse a prit peu d'efforts	1	2	3	4	5
J'ai obtenu les réponses que je cherchais sans trop d'échanges	1	2	3	4	5
Mes interactions avec le Chatbot ont été efficaces	1	2	3	4	5
Le montant d'information échangé entre moi et le Chatbot était adéquat	1	2	3	4	5

SVP continuez au verso



Tout à fait en désaccord
Plutôt en désaccord
Ni d'accord ni en désaccord
Plutôt d'accord
Tout à fait d'accord

Convivialité:

Ce Chatbot interagissait avec moi de manière respectueuse	1	2	3	4	5
Ce Chatbot interagit avec l'usage avec politesse	1	2	3	4	5
Mon interaction avec ce Chatbot a été amicale	1	2	3	4	5
Mes interactions avec ce Chatbot m'ont parfois fait sourire	1	2	3	4	5
Mes communications avec le Chatbot étaient conviviales	1	2	3	4	5

Communication personnalisée:

Je sentais comme une communication personnelle reçue de ce chatbot	1	2	3	4	5
Les informations fournies par le chatbot me paraissaient adaptées à mes besoins spécifiques	1	2	3	4	5
Le chatbot a ajusté ses réponses et recommandations à mes préférences	1	2	3	4	5

Innovativité:

Ce Chatbot est innovant	1	2	3	4	5
Ce Chatbot est unique	1	2	3	4	5
Ce Chatbot est technologiquement avancé	1	2	3	4	5

Apparence visuelle:

L'apparence du Chatbot est agréable	1	2	3	4	5
L'interface est attirante	1	2	3	4	5
Côté visuel, ce Chatbot est plaisant	1	2	3	4	5

Qualité globale:

Ce Chatbot est de bonne qualité	1	2	3	4	5
Ce Chatbot est bien conçu	1	2	3	4	5
<i>Je crois que ce Chatbot a besoin d'être amélioré (item renversé)</i>	1	2	3	4	5

Âge : ☐ 18 à 22ans ☐ 23 à 27ans ☐ 28 à 32ans ☐ 33 à 37ans
☐ 38 à 42ans ☐ 43 à 47ans ☐ 48 à 52ans ☐ 53 ans et +
☐ Préfère ne pas dire

Sexe : ☐ Homme ☐ Femme ☐ Préfère ne pas dire

Plus haut niveau de scolarité complété: ☐ Secondaire ☐ CEGEP ☐ Bachelier
☐ Maîtrise ☐ Doctorat
☐ Autre diplôme universitaire

Merci de votre collaboration!

ANNEXE G

CERTIFICAT D'APPROBATION ÉTHIQUE

CERTIFICAT D'APPROBATION ÉTHIQUE

Le Comité d'éthique de la recherche pour les projets étudiants impliquant des êtres humains (CERPE 3: sciences et sciences de l'éducation) a examiné le projet de recherche suivant et le juge conforme aux pratiques habituelles ainsi qu'aux normes établies par la *Politique No 54 sur l'éthique de la recherche avec des êtres humains* (Janvier 2016) de l'UQAM.

Titre du projet:	Identification et validation des facteurs qui influencent les perceptions de qualité d'un Chatbot
Nom de l'étudiant:	Seyf Eddine GUESMIA
Programme d'études:	Maîtrise en informatique de gestion (profil sans mémoire)
Direction de recherche:	Daniel TOMIUK

Modalités d'application

Toute modification au protocole de recherche en cours de même que tout événement ou renseignement pouvant affecter l'intégrité de la recherche doivent être communiqués rapidement au comité.

La suspension ou la cessation du protocole, temporaire ou définitive, doit être communiquée au comité dans les meilleurs délais.

Le présent certificat est valide pour une durée d'un an à partir de la date d'émission. Au terme de ce délai, un rapport d'avancement de projet doit être soumis au comité, en guise de rapport final si le projet est réalisé en moins d'un an, et en guise de rapport annuel pour le projet se poursuivant sur plus d'une année. Dans ce dernier cas, le rapport annuel permettra au comité de se prononcer sur le renouvellement du certificat d'approbation éthique.

Marie Nadeau R.N.

Marie Nadeau
Marie Nadeau
Présidente du CERPE 3 : Faculté des sciences, faculté des sciences de l'éducation
Professeure, Département didactique des langues

ANNEXE H

CERTIFICAT D'ACCOMPLISSEMENT
(EPTC 2 : FER)

Groupe en éthique
de la recherche

Piloter l'éthique de la recherche humaine

EPTC 2: FER

Certificat d'accomplissement

Ce document certifie que

Seyf Eddine Guesmia

*a complété le cours : l'Énoncé de politique des trois Conseils :
Éthique de la recherche avec des êtres humains :
Formation en éthique de la recherche (EPTC 2 : FER)*

24 novembre, 2017