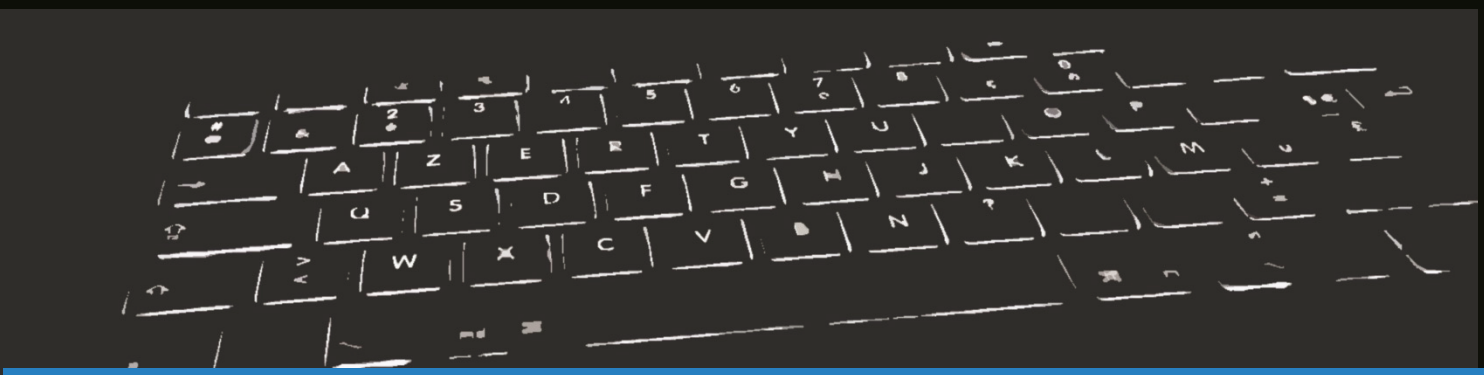
UNIVERSITE BRETAGNE LOIRE





THÈSE / Université Rennes 2

sous le sceau de l'Université européenne de Bretagne

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE RENNES 2

Mention : Psychologie

École doctorale : Éducation, Langage, Interactions, Cognition, Clinique (ED 603 ELICC)

présentée par

Nicolas MARTIN

Préparée au sein du LP3C (EA 1285)
Université Rennes 2
Laboratoire de Psychologie : Cognition, Comportement, Communication

Acceptabilité, Acceptation et Expérience Utilisateur : évaluation et modélisation des facteurs d'adoption des produits technologiques

Thèse soutenue le 26 Mars 2018
devant le jury composé de :

Aline CHEVALIER
Professeure des Universités, Université Toulouse Jean Jaurès / *Examinatrice*

Ludovic LE BIGOT
Professeur des Universités, Université de Poitiers / *Rapporteur*

André TRICOT
Professeur des Universités, Université Toulouse Jean Jaurès / *Rapporteur*

Éric JAMET
Professeur des Universités, Université Rennes 2 / *Directeur de thèse*





UNIVERSITÉ RENNES 2 – HAUTE BRETAGNE

Laboratoire de Psychologie : Cognition, Comportement, Communication (LP3C)

École Doctorale Education, Langage, Interactions, Cognition, Clinique
(ED 603 ELICC)

Sous le sceau de l'Université Bretagne Loire

Acceptabilité, Acceptation et Expérience Utilisateur : évaluation et modélisation des facteurs d'adoption des produits technologiques

Thèse de Doctorat

Discipline : Psychologie

Présentée par Nicolas MARTIN

Directeur de thèse : Éric JAMET

Soutenue le 26 Mars 2018

Jury :

Aline CHEVALIER
Professeure des Universités, Université Toulouse Jean Jaurès (*Examinatrice*)

Ludovic LE BIGOT
Professeur des Universités, Université de Poitiers (*Rapporteur*)

André TRICOT
Professeur des Universités, Université Toulouse Jean Jaurès (*Rapporteur*)

Eric JAMET
Professeur des Universités, Université Rennes 2 (*Directeur de thèse*)

Remerciements

Ce projet s'achève et il a été l'occasion pour moi de vivre de nombreuses expériences. Ainsi, je tiens à remercier toutes les personnes qui ont contribué à sa réussite de près ou de loin.

Tout d'abord, je voudrais remercier Éric Jamet. Merci pour toutes les étapes qui ont conduit à la soutenance de cette thèse. Merci de m'avoir donné ma chance en L3 en m'accordant un stage avec Sylvain au LOUSTIC puis pour toutes les autres étapes qui s'en sont suivi. Enfin, merci pour l'encadrement de cette thèse [C'était pas si mal une thèse sur autre chose que des docs d'apprentissage finalement, non ? ;)] !

Merci à Séverine Erhel et Géraldine Rouxel pour leurs rôles respectifs dans le suivi de cette thèse. Vos remarques et interventions auront toujours permis d'améliorer ce travail.

Je tiens à remercier Aline Chevalier, Ludovic Le Bigot et André Tricot d'avoir accepté de consacrer du temps à examiner cette thèse et de participer à ma soutenance.

Merci à tous les participants qui ont rempli mes questionnaires, je sais que parfois certains (pour ne pas dire la majorité) ont trouvé cela long (et n'ont pas manqué de le faire remarquer…) !

Cette thèse a été réalisée au sein d'une entreprise (pardon Lucie, une Fondation de Coopération Scientifique) remplie de brillantes personnes. D'abord, merci à b<>com (enfin le programme d'investissement d'avenir ANR-10-AIRT-07) d'avoir financé ma thèse et les belles expériences qui l'ont composée. Merci à toutes les personnes qui m'ont supporté et qui ont rendu mon quotidien vraiment agréable sur le plateau et plus largement pendant ces 3 dernières années. Je pense particulièrement à Alice, Amine, Anthony, Aurélie, Boris, Delphine, Karen, Lucie, Maud, Marion, Thomas, Véronique et tous les autres. PS : Pas de bol pour vous, j'ai re-signé pour 3 ans de plus ;) !

Je voudrais également remercier les membres de l'équipe UA dans laquelle j'ai travaillé et notamment Edouard, Florian, Grégoire, Guillaume, les Matthieu(s), Nico, Sonia.

Un remerciement vraiment spécial pour mes plus-que-collègues, la *dreamteam* Jean-Marc et Martin. Vous avez été toujours été là pour les joies (et les peines) et contribué largement à mon épanouissement au travail. Merci Martin pour tes nombreuses relectures, tes conseils et d'être mon inséparable binôme homonyme au bureau !

Je remercie aussi les collègues et bien plus du labo CRPCC/LP3C, qui m'ont accueilli lorsque j'étais encore un tout jeune étudiant : Anne-Laure, Barbara, Jean, Jérémie, Jonathan, Martin, Natacha, Sylvain. Vous m'avez toujours accompagné tant personnellement que professionnellement, donné le goût de la recherche et permis d'affuter mes capacités de débat. Mais c'est de votre faute si maintenant je vous parle de stats dès que j'en ai l'occasion ;) ! Mention spéciale à Sylvain pour avoir été mon modèle de doctorant et pour m'avoir supporté comme stagiaire pendant presque 1 an ! Un merci aussi au reste du labo, je pense notamment à Benjamin, Emeline, Maxime, Julie, Tiphaine.

Il y a aussi les remerciements inclassables, je pense notamment à Madeleine LeGarff-Romé et Frédéric Mathieu qui ont contribué à leur façon à mon parcours de vie.

Merci aux membres de ma très grande famille et de celle de Salomé qui m'ont soutenu de près ou de loin.

Merci aussi aux amis qui m'ont aidé sur le plan personnel ou professionnel à toujours aller de l'avant et qui ont permis que tout ceci arrive à son terme. Je pense particulièrement à Arnaud, Cyrille, Gaëtan, Mathilde, Marie, Maximilien, Mélissa, Paul-Baptiste, Pierick, Pierre-Maël, Salma.

Une mention spéciale pour Anna : c'est quand même un peu grâce à toi si j'ai osé prendre cette voie si spéciale de la recherche. Merci pour tes soutiens, tes conseils, et tout le reste.

Enfin, un grand merci pour la meilleure, ma Salomé. Merci d'être là tous les jours à mes côtés pour m'écouter, me conseiller et me supporter. Ces dernières années ont été particulièrement riches de joies, d'expériences et de moments mémorables. Merci d'être ce que tu es et d'embellir mon quotidien.

“Essentially, all models are wrong, but some are useful”.

(George E. P. Box, 1987)

Préambule

Les travaux de cette thèse ont fait l'objet de plusieurs publications acceptées ou soumises :

Martin, N., Jamet, É., Erhel, S., & Rouxel, G. (Soumis). Which factors influence behavioral intention before use? Studies with two mobile applications using structural equation modeling.

Martin, N., Jamet, É., Erhel, S., & Rouxel, G. (2016). Evaluation of technological products in mobility context: first steps toward new scales with single items. In *Proceedings of the 28th Conference on l'Interaction Homme-Machine* (pp. 237-245). ACM Press. https://doi.org/10.1145/3004107.3004133

Martin, N., Jamet, É., Erhel, S., & Rouxel, G. (2016). From Acceptability to Acceptance: Does Experience with the Product Influence User Initial Representations? In C. Stephanidis (Ed.), *HCI International 2016 – Posters' Extended Abstracts* (Vol. 617, pp. 128–133). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40548-3_21

Martin, N., Erhel, S., Jamet, É., & Rouxel, G. (2015). Quels liens entre expérience utilisateur et acceptabilité ? In *Proceedings of the 27th Conference on l'Interaction Homme-Machine* (pp. 34-39). ACM Press. http://doi.org/10.1145/2820619.2825015

Table des matières

Remerciements **3**

Préambule **6**

Table des matières **7**

Introduction théorique **11**

1. Contexte **11**

1.1. Projet UXPPE 11

1.2. Termes et dénominations 12

2. Introduction générale **13**

3. Modèles de l'acceptabilité et de l'acceptation **15**

3.1. TAM 16

3.2. UTAUT 19

3.3. Mesures de l'acceptabilité et de l'acceptation 24

4. Les déterminants affectivo-motivationnels **26**

4.1. Les émotions 26

4.2. Motivation, plaisir perçu et fun 30

4.3. Engagement 32

4.4. Le *Flow* 35

4.5. Confiance, fiabilité et crédibilité 37

4.6. Superposition des variables affectivo-motivationnelles 38

5. Modèles du jugement utilisateur et de l'expérience utilisateur (UX) **39**

5.1. La notion d'Utilisabilité 39

5.2. La notion d'expérience utilisateur (UX) 40

5.3. Qualité d'expérience 50

5.4. Outils de mesure de l'UX 52

5.5. Synthèse des modèles de l'UX 54

6. Modèles intégrateurs des notions UX et d'acceptabilité/acceptation **56**

6.1. Modèle de van Schaik et collaborateurs 56

6.2. Modèle de Février 58

7. Évolution du jugement....61
7.1. Évolution au cours du temps....61
7.2. Outils de mesure de l'évolution au cours du temps....63
8. Approche proposée....65
8.1. Les éléments essentiels....65
8.2. Modèle théorique intégratif....66
9. Conclusion....68
9.1. Une variété de modèles existants....68
9.2. Problématique de la thèse....69

Chapitre 1 : Modélisation de l'adoption avant usage....70

1. Acceptabilité / Acceptation et expérience utilisateur : jugement avant et après usage....71
2. Rappel du contexte théorique et hypothèses....72
2.1. Acceptabilité et Acceptation....72
2.2. Expérience utilisateur (UX)....75
3. Étude 1 : Application Sleep Better....78
3.1. Méthode....78
3.2. Résultats....81
3.3. Conclusion....88
4. Étude 2 : Application Google Fit....89
4.1. Méthode....89
4.2. Résultats....90
5. Discussion....96
6. Limites et futurs travaux....98
7. Conclusion....98

Chapitre 2 : Jugement après usage....100

1. Introduction....100
2. Les variables influençant le jugement après usage....100
2.1. Les qualités fonctionnelles....101
2.2. Les facteurs affectivo-motivationnels....102
3. Étude : application Waze....104
3.1. Méthode....104
3.2. Résultats....107
3.3. Discussion....114

4. Conclusion....117

Chapitre 3 : Évolution du jugement utilisateur au cours du temps.... 118

1. Introduction....118

1.1. Théorie des Attentes-Infirmations (EDT - *Expectation-Disconfirmation Theory*) 119

1.2. EDT et adoption des produits technologiques....119

2. Étude....120

2.1. Méthode....121

3. Résultats....123

4. Discussion et conclusion....126

Chapitre 4 : Construction des échelles à items uniques pour la mesure avant et après usage.... 128

1. Introduction....128

2. Les échelles à items uniques....128

3. Principes statistiques sous-jacents aux échelles à items uniques....129

4. Création des échelles de mesure avant et après usage à items uniques....133

4.1. Échelles avant usage à items uniques....134

4.2. Échelles après usage à items uniques....138

5. Discussion et Conclusion....140

Chapitre 5 : Preuve de concept - *Valeurs seuils et librairie R*.... 144

1. Création des valeurs seuils pour les questionnaires....144

1.1. Régression logistique....145

1.2. Expérimentation....146

2. Preuve de concept : librairie R....153

Chapitre 6 : Discussion Générale.... 155

1. Acceptabilité, UX et variables affectivo-motivationnelles....155

1.1. Résumé des fondements théoriques....155

1.2. Principaux résultats....157

2. Échelles de mesure....159

3. Pistes d'améliorations et futurs travaux....160

3.1. Produits évalués et conditions expérimentales....161

3.2. Échantillons et méthodes de recueils des données....162

3.3. Modèles proposés....164

3.4. Mesure du jugement utilisateur et de l'adoption 165
3.5. Temporalité 166

Conclusion 168

Références 169

Annexes 204

Index des tableaux 252

Index des figures 255

Introduction théorique

1. Contexte

Avant d'exposer en détail les modèles rattachés aux champs théoriques de l'acceptabilité et de l'expérience utilisateur (UX), il semble important de présenter le contexte et le projet dans lequel cette thèse a été réalisée.

1.1. Projet UXPPE

Cette thèse a été réalisée au sein de l'Institut de Recherche Technologique (IRT) b<>com de Rennes soutenu par le gouvernement français à travers le Programme d'Investissement d'Avenir (PIA) portant la référence ANR-A0-AIRT-07. Plus précisément, cette thèse a été effectuée au sein du laboratoire Usages et Acceptabilité (UA) dans le cadre du projet UXPPE (*User eXperience, ProsPective and Evaluation*). Ce projet visait à apporter une dimension de recherche et d'innovation sur les méthodologies d'analyses, de modélisations, et d'évaluations de l'expérience utilisateur dans les technologies de l'image, de l'interaction et de l'immersion ainsi qu'une prospective sur les usages des futurs services immersifs. La thèse était intégrée au sein d'un des lots du projet (lot 4.1.) ayant pour objectif principal le développement d'un nouvel outil pour l'évaluation de produits technologiques par les utilisateurs. Cet outil devait répondre à une double demande : être à la fois simple d'usage pour les participants et les expérimentateurs, tout en permettant d'évaluer le plus exhaustivement possible les dimensions rattachées à l'adoption d'un produit technologique.

Cette volonté de construire de nouveaux outils de mesure repose sur la présence concomitante des champs théoriques de l'acceptabilité (e.g., Venkatesh, Morris, Davis, & Davis, 2003) et de l'expérience utilisateur (e.g., Hassenzahl, 2004) apparaissant comme potentiellement complémentaires mais que très rarement mis en lien dans la littérature. Ainsi, il était tout d'abord nécessaire de discerner les contours de ces différentes théories et d'extraire les variables semblant être les plus pertinentes. En conséquence, l'état de l'art vise à exposer les théories, modèles et variables les

plus courants pour comprendre et évaluer le jugement envers un produit technologique.

1.2. Termes et dénominations

Au sein de cette thèse, plusieurs termes ont été choisis pour nommer certains concepts et nécessitent d'être explicités avant d'exposer davantage les modèles théoriques associés. Tout d'abord, les termes **jugement utilisateur** et **adoption** seront employés pour désigner l'évaluation qu'un individu émet envers un produit technologique. Ces expressions ont été choisies afin d'éviter d'emprunter les désignations spécifiques aux champs théoriques de l'acceptabilité et de l'expérience utilisateur. En effet, le terme acceptabilité est largement utilisé lorsque l'on s'intéresse au consentement ou au refus d'un individu à utiliser un système. Cependant, une large partie de la thèse est dédiée aux modèles théoriques de l'acceptabilité. Afin de ne pas mélanger le fait qu'un individu soit prêt à utiliser un système et le champ théorique de l'acceptabilité, le terme **adoption** sera utilisé lorsqu'on s'intéresse au jugement et l'appréciation que les individus émettent envers un comportement d'usage d'un produit. Le **jugement utilisateur** se réfère, quant à lui, à un élément plus large que l'adoption d'un produit. Il s'agit de l'avis global émis par l'utilisateur envers un produit technologique. L'évaluation des qualités du produit mais également à l'évaluation globale envers un système y sont liées.

Enfin, les termes **système**, **service** et **produit** seront utilisés indifféremment. Ils désignent les technologies informatiques utilisables par les individus : cela concerne, par exemple, les ordinateurs, les applications mobiles et les logiciels.

2. Introduction générale

Les produits technologiques font partie, un peu plus chaque jour, de notre quotidien. De ce fait, ils jouent un rôle de plus en plus important sur les comportements humains et les systèmes économiques (OCDE, 2003). Cette large influence a conduit au développement d'un important domaine de recherche (eh & Teng, 2012) autour de l'exploration des facteurs centraux de l'évaluation et de l'adoption d'un produit par les utilisateurs. Historiquement, ces travaux ont été conduits autour des concepts d'acceptabilité et d'acceptation d'un produit (e.g., Venkatesh et al., 2003). Ils poursuivent l'objectif d'extraire les variables clés dans l'adoption d'un produit (Dubois & Bobillier-Chaumon, 2010). Ces travaux sont majoritairement orientés autour de l'évaluation des facteurs rattachés aux qualités fonctionnelles des produits. Plus récemment, afin de tenir compte des travaux mettant en avant l'influence de facteurs non fonctionnels (e.g., Harbich & Hassenzahl, 2016), un nouveau champ de recherche autour de la notion d'expérience utilisateur (UX) a fait son apparition. L'UX vise à englober à la fois les composantes fonctionnelles des produits mais également les qualités non fonctionnelles (e.g., l'esthétisme) ayant une influence sur l'expérience avec un produit (Hassenzahl, 2004). Cependant, comme il sera présenté par la suite, les définitions de ces concepts ne font pas encore l'unanimité au sein de la littérature (Lallemand, Gronier, & Koenig, 2015). Il semble ainsi exister une pluralité de conceptions autour de ces dimensions. Au-delà des approches théoriques autour de l'acceptabilité et de l'UX, des facteurs rattachés au fonctionnement humain, que nous appellerons variables affectivo-motivationnelles, ont fait l'objet de nombreux travaux (e.g., motivation intrinsèque et extrinsèque ; F. D. Davis, Bagozzi, & Warshaw, 1992). L'introduction de ces variables vise à pallier une des limites des modèles théoriques de l'acceptabilité (e.g., TAM). En effet, il est généralement présupposé, au sein de ces modèles, que le jugement envers un système se construit sur une évaluation rationnelle des critères de qualité d'un produit (Février, Jamet, & Rouxel, 2012). Le fonctionnement humain est plus complexe, nos réactions affectivo-motivationnelles influencent notamment nos comportements ; par conséquent en tenir compte pourrait permettre de mieux comprendre les mécanismes du ressenti utilisateur (Saariluomaand & Jokinen, 2014).

Afin de proposer une vue d'ensemble des avancées respectives de ces champs théoriques, une synthèse des travaux scientifiques et notamment des modèles

théoriques développés au sein de ces divers domaines sera proposée. Certaines variables se superposent sur le plan conceptuel au point parfois d'être mesurées de manière similaire (par exemple, le plaisir perçu est généralement évalué avec des items de motivation intrinsèque ; Balog & Pribeanu, 2010). Par conséquent, il semble essentiel de détailler ces différentes variables afin de sélectionner les plus pertinentes. Dans ce chapitre, une première partie sera dédiée à la présentation des principaux modèles de l'acceptabilité et de l'acceptation. Puis, un ensemble de variables affectivo-motivationnelles apparues au cours des dernières années pour compléter notamment les différents modèles de l'acceptabilité sera présenté. Enfin, une revue des modèles de l'UX, plus récents, sera proposée. A la suite de la présentation des modèles théoriques, un panorama des outils de mesure rattachés sera proposé.

3. Modèles de l'acceptabilité et de l'acceptation

Depuis de nombreuses années, des travaux scientifiques ont porté sur les facteurs qui conduisent les individus à adopter tels ou tels produits (voir Khechine, Lakhal, & Ndjambou, 2016 ; King & He, 2006 pour une revue des travaux sur les principaux modèles de l'acceptabilité). Ces travaux ont été appliqués à des systèmes variés comme des Environnements Informatiques pour l'Apprentissage Humain (e.g., Tricot et al., 2003) ou des jeux vidéo (e.g., Hsu & Lu, 2004). Ainsi, les notions d'acceptabilité et d'acceptation (Dubois & Bobillier-Chaumon, 2010) ont été explorées par de nombreux chercheurs et les principaux résultats seront présentés dans la partie ci-dessous. L'accent sera mis sur les modèles TAM (F. D. Davis, 1989) et UTAUT (Venkatesh et al., 2003 ; Venkatesh, Thong, & Xu, 2012a) qui semblent les plus reconnus au sein de la littérature et les plus pertinents pour ces travaux de thèse.

Il semble, tout d'abord, nécessaire de distinguer les notions d'acceptabilité et d'acceptation même si elles sont couramment employées comme des synonymes dans la littérature. Ces notions renvoient à des temporalités différentes dans l'évaluation (Adell, Varhelyi, & Nilsson, 2014 ; Bobillier-Chaumon & Dubois, 2010). D'une part, l'acceptabilité correspondrait à une forme de jugement prospectif. Elle peut être définie comme la tendance à être plus ou moins favorable à l'usage d'une technologie avant sa diffusion (Schade & Schlag, 2003 ; Schuitema, Steg, & Forward, 2010). D'autre part, on retrouve la notion d'acceptation. Elle est définie comme étant le jugement et les réactions comportementales envers un produit après usage (Bobillier-Chaumon & Dubois, 2010 ; Schuitema et al., 2010). Ainsi, l'acceptation ne peut être évaluée qu'après l'utilisation du système et correspondrait à l'usage d'un système pour les tâches pour lesquelles il a été conçu (Dillon & Morris, 1996). Les notions d'acceptabilité et d'acceptation peuvent être vues comme faisant partie d'un continuum commençant avant usage et allant jusqu'à l'usage prolongé (Bobillier-Chaumon & Dubois, 2010). Malgré cette distinction entre l'évaluation avant et l'après usage mise en avant dans la littérature, la plupart des travaux sont réalisés sur des données collectées après au moins une session d'usage (Karahanna, Straub, & Chervany, 1999 ; Šumak & Šorgo, 2016 ; S. ang, Lu, Gupta, Cao, & Zhang, 2012). De plus, ces travaux autour de l'acceptabilité et de l'acceptation (e.g., L. Chen,

Meservy, & Gillenson, 2012) ont en commun l'intention d'usage déclarée, comme mesure pour rendre compte de l'utilisation d'une technologie particulière par un individu (Sezgin & ıldırım, 2016).

Ainsi, un premier modèle, le TAM (F. D. Davis, 1989), s'est répandu dans le champ théorique de l'acceptation.

3.1. TAM

Le TAM, modèle développé par Davis (1989), vise à expliquer les facteurs de l'adoption des produits technologiques. Ce modèle a connu un essor considérable depuis sa création et apparait aujourd'hui comme le plus connu (Sorgenfrei, Ebner, Smolnik, & Jennex, 2014). Celui-ci repose sur les théories de l'action raisonnée (Fishbein & Ajzen, 1975) et du comportement planifié (Ajzen, 1991) qui visent à prédire le comportement humain.

Selon la Théorie de l'Action Raisonnée (*Theory of Reasoned Action* - TRA ; Fishbein & Ajzen, 1975), le comportement d'un individu (e.g., utiliser un système) est déterminé par l'intention de la personne à réaliser ce comportement. Cette intention serait influencée par les attitudes de l'individu (déterminées par la perception de l'individu des conséquences de ce comportement et par l'évaluation de ces conséquences ; Fishbein & Ajzen, 1975, p. 301) et ses normes subjectives (i.e., « la perception de l'individu sur le fait que la plupart des personnes qui sont importantes à ses yeux pensent qu'il devrait ou ne devrait pas effectuer un comportement » [notre traduction] ; Fishbein & Ajzen, 1975, p. 302). Le modèle complet du TRA est présenté en Figure 1.

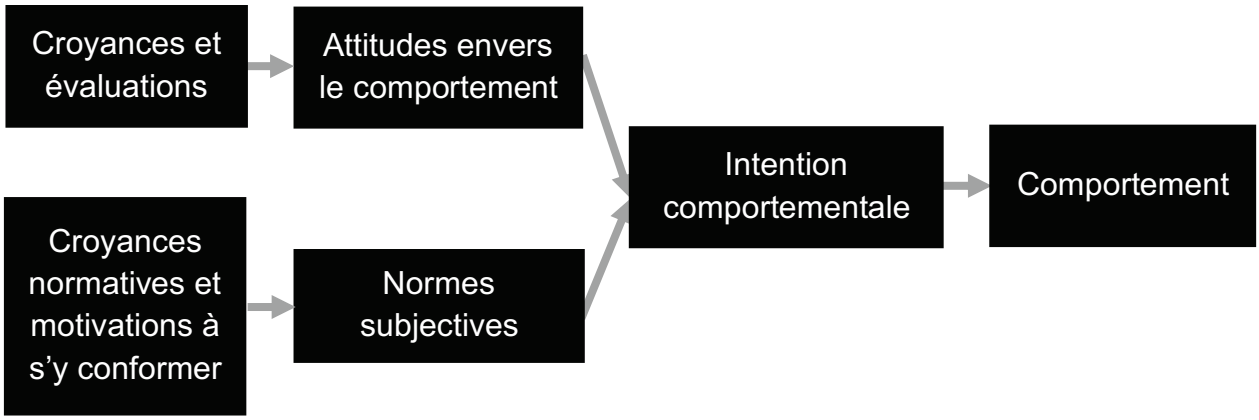


Figure 1. Théorie de l'Action Raisonnée adaptée de Davis, Bagozzi et Warshaw (1989)

Selon la Théorie du Comportement Planifié (*Theory of Planned Behavior* - TPB ; Ajzen, 1991), l'intention comportementale (i.e., la force de l'intention d'exercer un comportement spécifique ; Fishbein & Ajzen, 1975) reposerait sur plusieurs éléments : la capacité de l'individu à réaliser un comportement, la présence de ressources suffisantes pour réaliser ce comportement, les possibilités d'atteindre les résultats attendus et la perception par l'individu de ces possibilités. Le modèle complet du TPB est présenté en Figure 2.

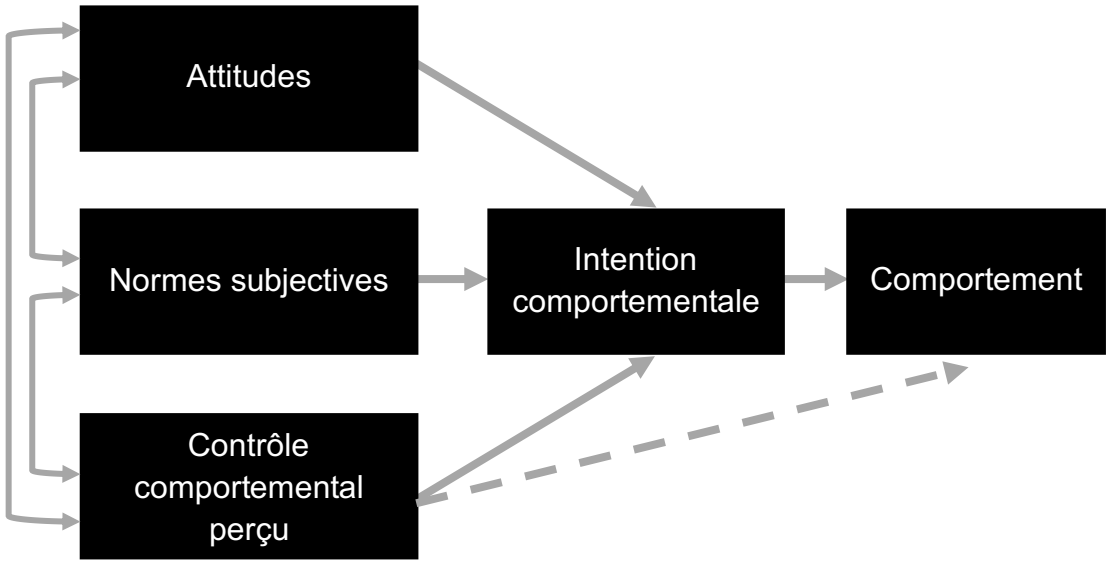


Figure 2. Théorie du Comportement Planifié adaptée de Ajzen (1991)

Le TAM a ainsi été construit sur la base des travaux de la TRA et TPB. Selon le TAM (voir figure 3), les individus percevraient les systèmes selon deux propriétés : la facilité d'usage perçue et l'utilité perçue. La facilité d'usage perçue renvoie au « degré selon lequel un individu croit qu'utiliser un système lui demandera peu ou pas effort » [notre traduction] (F. D. Davis, 1989, p. 320). Quant à l'utilité perçue, elle renvoie au « degré selon lequel un individu croit qu'un système peut améliorer sa performance » [notre traduction] (F. D. Davis, 1989, p. 320). Ces deux facteurs influenceraient les attitudes envers l'usage, qui à leur tour joueraient un rôle sur l'intention d'usage[1]. L'intention d'usage, quant à elle, prédirait l'usage réel.

[1] L'intention d'usage est la variable commune aux travaux sur l'acceptabilité pour évaluer l'usage. Elle renvoie à l'usage d'un système particulier par des individus (Sezgin & Yıldırım, 2016).

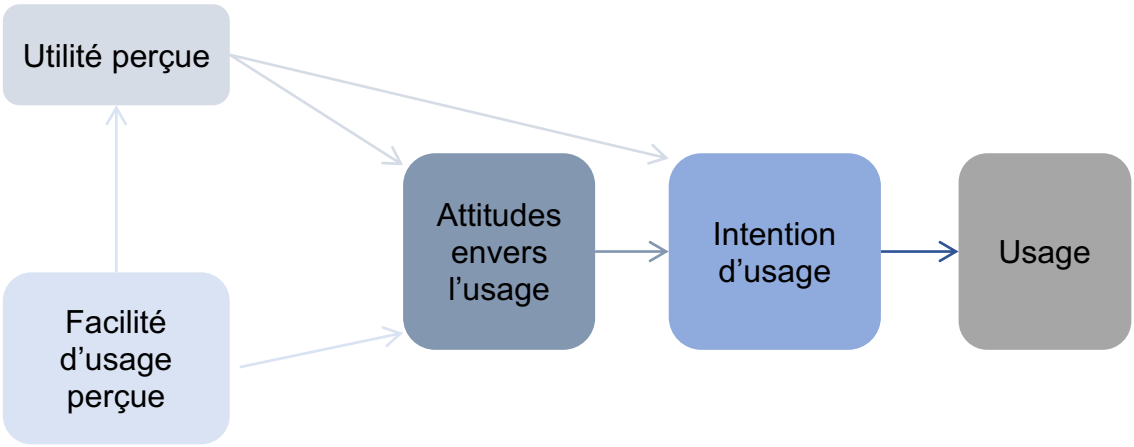


Figure 3. TAM adapté de Davis (1989)

Ce modèle a servi de base à de nombreux travaux (e.g., Ortega Egea & Román González, 2011 ; Schepers & Wetzels, 2007) ayant fait l'objet de plusieurs méta-analyses (King & He, 2006 ; Ma & Liu, 2004 ; Schepers & Wetzels, 2007). Selon ces diverses études (voir Tableau 1), l'utilité perçue jouerait un rôle majeur dans l'intention d'usage.

Tableau 1. Influence de l'Utilité et de l'Utilisabilité sur l'Intention d'Usage selon 4 méta-analyses

	Utilité perçue → Intention d'usage	**Facilité d'Usage perçue → Intention d'usage**
Dwivedi et al. (2011)	.34	.14
King & He (2006)	.51	.19
Ma & Liu (2004)	.41	.28
Schepers & Wetzels (2007)	.56	.41
Moyenne	*.46*	*.26*

La popularité du TAM provient probablement de sa facilité d'application et d'interprétation, grâce notamment au questionnaire proposé (voir Annexe 1) pour évaluer la facilité d'usage perçue et l'utilité perçue (Venkatesh, Davis, & Morris, 2007).

En dépit de sa large diffusion, ce modèle présente un certain nombre de limites. Tout d'abord, son pouvoir explicatif de l'intention d'usage est relativement limité (Legris, Ingham, & Collerette, 2003 ; Taylor & Todd, 1995b) avec une variance expliquée de l'intention comportementale en moyenne de 40 % (Legris et al., 2003 ; Taylor & Todd, 1995b). Le TAM est également critiqué pour la variabilité retrouvée à

travers les études sur les liens entre les variables. En effet, même si les corrélations entre les variables sont généralement significatives et relativement fortes, King et He (2006) notent, dans leur méta-analyse, des écarts non négligeables entre les études qui pourraient notamment s'expliquer par la présence de variables modératrices non prises en compte dans le modèle (King & He, 2006).

Deuxièmement, ce modèle reposerait sur une vision trop déterministe et rationnelle de l'acceptation (Brangier, Hammes-Adelé, & Bastien, 2010). En effet, au sein du TAM, il est supposé que l'individu prend une décision rationnelle uniquement sur la base d'une perception d'utilité et de facilité d'usage du produit. Ainsi, la prise en compte de variables dites sociales apparait nécessaire pour améliorer ce modèle (Brangier et al., 2010).

Face à ces critiques, différents chercheurs ont tenté d'inclure de nouvelles variables afin d'augmenter le pouvoir explicatif du modèle sur l'intention d'usage (e.g., S. Ha & Stoel, 2009 ; Shih, 2004), mais la variance expliquée reste limitée (Brangier, Dufresne, & Hammes-Adelé, 2009). Ainsi, d'autres modèles ont donc vu le jour comme l'UTAUT (2003), présenté ci-dessous, qui vise à proposer la synthèse des principaux modèles de l'acceptation.

3.2. UTAUT

Venkatesh, Morris, Davis et Davis (2003) ont développé l'UTAUT dans le but d'obtenir un meilleur pouvoir explicatif de l'intention d'usage au travers d'un modèle unifié. Pour construire ce modèle, ils ont tout d'abord réalisé une revue des huit modèles théoriques suivants :

- La théorie de l'Action Raisonnée – TAR (Fishbein & Ajzen, 1975) ;
- Le TAM (F. D. Davis et al., 1989) ;
- Un modèle de la Motivation – MM (reposant sur un découpage entre motivation intrinsèque et extrinsèque) (F. D. Davis et al., 1992);
- La théorie du Comportement Planifié – TPB (Ajzen, 1991);
- La combinaison du TAM et de la théorie du comportement planifié – C-TAM-TPB (Taylor & Todd, 1995a) ;

- Le modèle de l'utilisation de l'ordinateur – MPCU (Thompson, Higgins, & Howell, 1991) ;
- La théorie de la diffusion de l'innovation – IDT (Rogers, 1995) ;
- La théorie sociale cognitive – SCT (Compeau & Higgins, 1995).

Afin de sélectionner les variables pertinentes pour la construction de leur nouveau modèle, quatre études longitudinales ont été réalisées (au sein de quatre entreprises différentes) à partir d'un questionnaire composé de 77 items issus de ces huit modèles théoriques (e.g., cet item est proposé pour mesurer la dimension utilité perçue du TAM « utiliser ce système me permettrait d'augmenter ma productivité » [notre traduction]). À l'aide d'une modélisation PLS (*Partial Least Squares*), les variables ayant une influence significative sur l'intention d'usage ont été extraites. Ainsi, sept construits sont ressortis comme des déterminants significatifs de l'intention d'usage. Ces derniers ont alors été regroupés au sein de quatre facteurs (voir figure 4) :

- *Attentes en termes d'effort* : degré de facilité associé à l'usage du système.
- *Attentes en termes de performance* : degré selon lequel un individu croit que l'utilisation du système l'aidera à atteindre un but.
- *Influence sociale* : degré selon lequel un individu perçoit que les personnes qui lui sont importantes croient qu'il ou elle devrait utiliser le nouveau système
- *Conditions facilitatrices* : degré selon lequel un individu perçoit qu'une organisation et un support existent pour l'aider dans l'utilisation du système.

De plus, quatre variables modératrices ont été ajoutées : l'âge, le sexe, l'expérience et le volontarisme à l'usage.

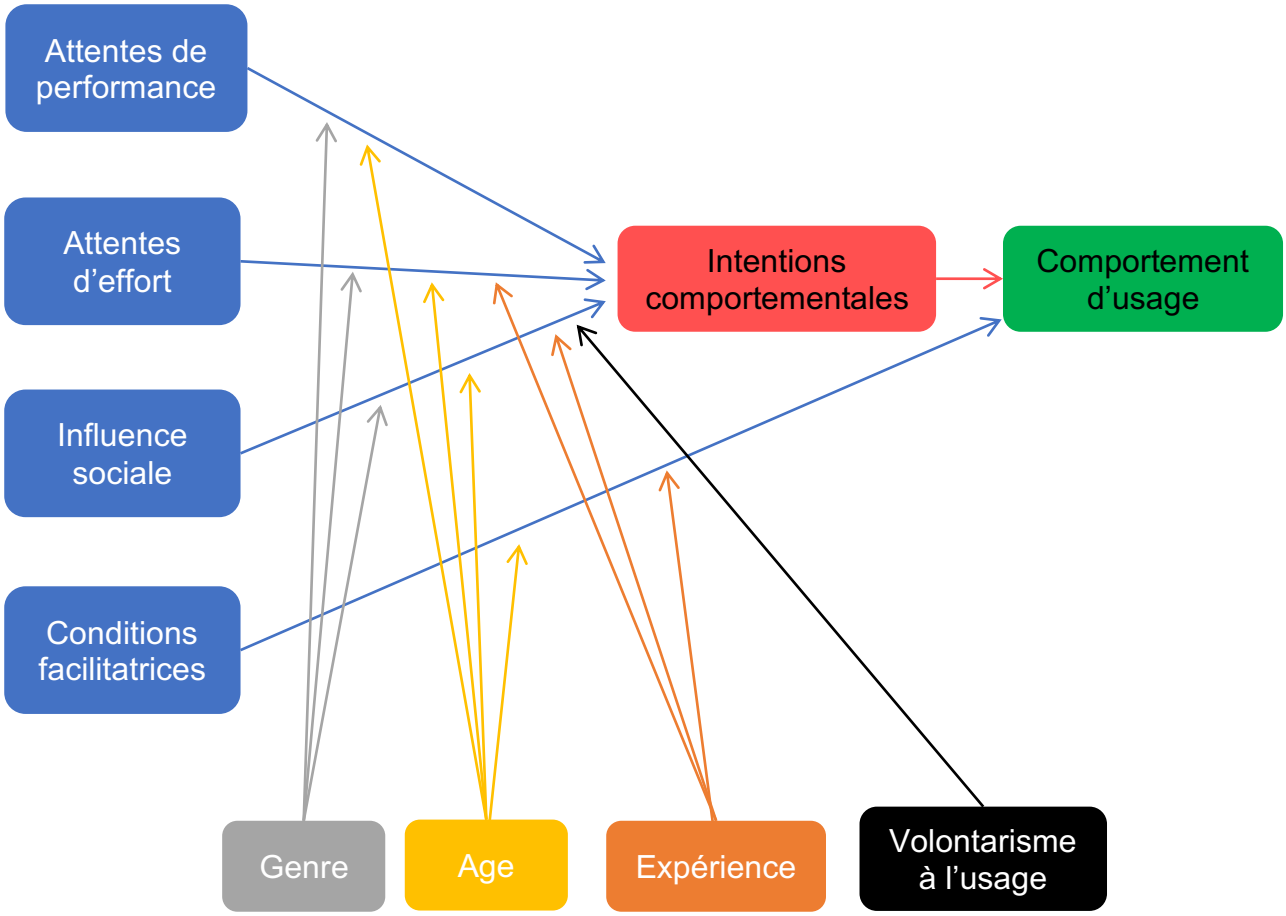


Figure 4. UTAUT (Venkatesh et al., 2003)

Venkatesh et ses collaborateurs (2003) ont conduit un examen du modèle de mesure avant de tester les liens entre les variables. Ils ont constaté que tous les items avaient des saturations supérieures à .70 dans les dimensions respectives dont ils dépendaient a priori. Ce résultat signifie que l'ensemble des items offre des qualités psychométriques satisfaisantes.

Au regard du nombre de paramètres à estimer et de la taille relativement modeste de leurs échantillons, Venkatesh et ses collaborateurs ont choisi de réduire le nombre d'items par facteur à quatre avant de tester leur modèle structurel. Ces items correspondent au questionnaire proposé par les auteurs pour évaluer les produits technologiques (voir Annexe 2). Leurs résultats ont conduit au modèle présenté en figure 4. Enfin, Venkatesh et ses collègues (2003) parviennent, à travers l'UTAUT, à expliquer environ 70 % de la variance de l'intention d'usage dans leurs différents échantillons.

Que cela soit dans l'article d'origine (Venkatesh et al., 2003) ou la méta-analyse de Dwivedi et ses collaborateurs (2011), il est constaté que ce sont les effets des attentes en termes de performance qui sont les plus souvent répliqués (voir tableau 2) et qui contribuent le plus à l'explication de l'intention d'usage (voir figure 5), comme pour le TAM (avec la variable utilité perçue).

Tableau 2. Pourcentages de relations significatives à travers différentes études issues d'une méta-analyse sur l'UTAUT (Dwivedi et al., 2011)[2]

Liens entre variables	**AP - IC**	**AE - IC**	**IS - IC**	**CF - IC**	**CF - U**	**IC - U**
Liens significatifs	25	19	22	9	14	8
Liens non significatifs	0	5	3	2	2	0
Total	25	24	25	11	16	8
Pourcentage	100%	79%	88%	82%	88%	100%

[2] AP : Attentes de Performance – AE : Attentes d'Effort – IS : Influence Sociale – CF : Conditions Facilitatrices – IC : Intention comportementale – U : Usage.

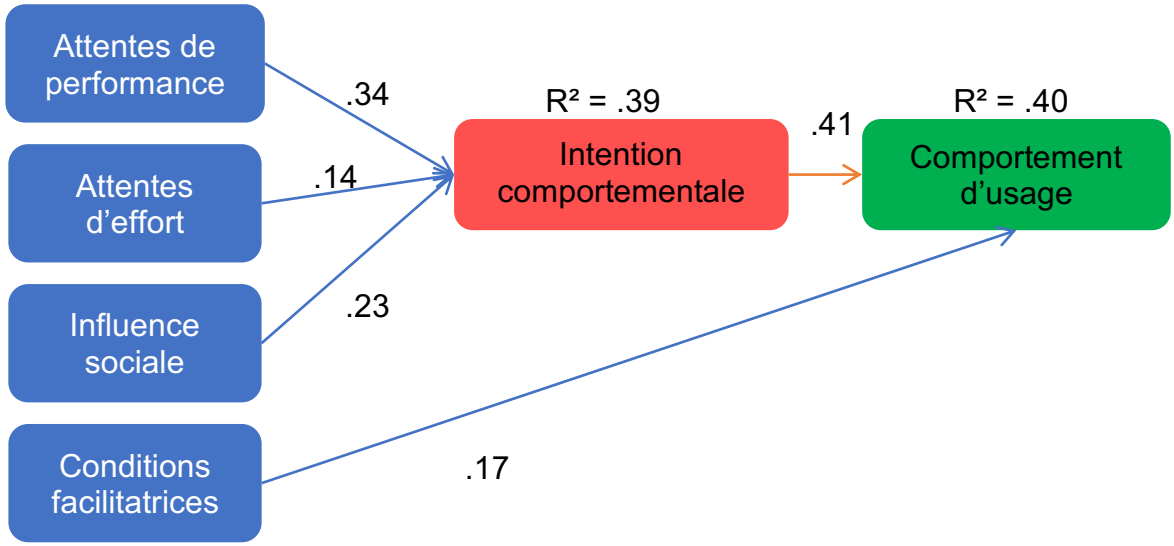


Figure 5. Coefficients bêta moyens des variables issus de la méta-analyse sur l'UTAUT (Dwivedi et al., 2011)[3]

Ainsi, selon la méta-analyse de Dwivedi et al. (2011) et la revue d'Attuquayefio et Addo (2014), à l'exception de quelques résultats inconsistants, la majorité des études rapporte des résultats en accord avec le modèle d'origine. Cependant, il est important de noter que les variables modératrices ne sont presque jamais prises en compte dans les études et par conséquent dans les méta-analyses (e.g., Dwivedi et al., 2011 ; Khechine et al., 2016).

Enfin, une nouvelle version de l'UTAUT a récemment été proposée (Venkatesh et al., 2012a) afin d'améliorer la prédiction du modèle et prendre en compte de nouvelles variables découvertes dans la littérature depuis la création de l'UTAUT. Ainsi, la motivation hédonique (« le fun ou le plaisir dû à l'utilisation d'une technologie » [notre traduction] ; Venkatesh et al., 2012a, p. 161), la valeur de prix (« le compromis cognitif des consommateurs entre les avantages perçus des applications et le coût monétaire pour les utiliser » [notre traduction] ; Dodds, Monroe, & Grewal, 1991) et les habitudes (« le construit perceptif qui reflète les résultats d'expériences antérieures » [notre traduction] ; Venkatesh et al., 2012a) ont été ajoutées comme prédicteurs de

[3] R^2 : coefficient de détermination. Ce chiffre entre 0 et 1 permet de déterminer à quel point les prédicteurs permettent d'expliquer la variable à prédire. Par exemple, dans ce modèle la variance de la variable « Comportement d'usage » est expliquée à 40 % (R^2 = .40) par « les conditions facilitatrices » et « l'intention comportementale ».

l'acceptabilité. Sur leurs données, ils parviennent à expliquer 74 % de la variance avec leur nouveau modèle au lieu de 56 % en utilisant la première version de l'UTAUT (Venkatesh et al., 2003).

3.3. Mesures de l'acceptabilité et de l'acceptation

Afin de construire et tester les modèles théoriques de l'acceptabilité et de l'acceptation, des outils de mesures ont été construits. Ainsi, les deux principaux modèles, le TAM (F. D. Davis, 1989) et l'UTAUT (Venkatesh et al., 2003, 2012a), ont été étudiés à travers de questionnaires spécialement développés. Ces échelles de mesure visent à évaluer, à l'aide de plusieurs items, chacune des dimensions issues de la littérature.

Le questionnaire original du TAM (F. D. Davis, 1989) comporte uniquement deux facteurs : l'utilité perçue (6 items, e.g., « utiliser [ce système] dans mon travail devrait me permettre d'accomplir mes tâches plus rapidement » [notre traduction]) et l'utilisabilité perçue (6 items, e.g., « apprendre à utiliser [ce système] devrait être facile pour moi » [notre traduction]).

En ce qui concerne l'UTAUT2 (Venkatesh et al., 2012a), le questionnaire est composé de mesures pour les attentes en termes d'effort (3 items, e.g., « Je trouve [ce système] utile dans ma vie quotidienne » [notre traduction]), les attentes en termes de performance (4 items, e.g., « Je trouve [ce système] facile à utiliser » [notre traduction]), l'influence sociale (3 items, e.g., « Les personnes dont j'apprécie l'opinion préfèrent que j'utilise [ce système] » [notre traduction]), les conditions facilitatrices (4 items, e.g., « J'ai les ressources nécessaires pour utiliser [ce système] » [notre traduction]), la motivation hédonique (3 items, e.g., « Utiliser [ce système] est fun » [notre traduction]), la valeur de prix (3 items, e.g., « Le prix de [ce système] est raisonnable » [notre traduction]), les habitudes (3 items, e.g., « L'utilisation de [ce système] est devenue une habitude pour moi » [notre traduction]) et l'intention d'usage (3 items, e.g., « J'ai l'intention de continuer à utiliser [ce système] dans le futur » [notre traduction]).

Malgré la popularité de ces modèles, ils font l'objet de critiques à propos de la vision trop rationnelle du fonctionnement humain qu'ils véhiculent. En effet, au sein de ces modèles, il est supposé que les individus conduisent une évaluation rationnelle des qualités du produit et qu'ils émettent leur jugement envers le système à partir de

ces évaluations. Or, le fonctionnement humain est plus complexe, ainsi de nombreux chercheurs ont tenté d'introduire des variables dites affectivo-motivationnelles aux modèles de l'acceptabilité pour en améliorer la précision. Un aperçu de ces travaux sera présenté dans le chapitre suivant.

En Résumé

Les théories de l'acceptabilité et de l'acceptation visent à découvrir les facteurs d'adoption ou de rejet des produits. De nombreux travaux ont ainsi été réalisés et ont abouti à deux modèles majeurs : le TAM (F. D. Davis, 1989) et l'UTAUT (Venkatesh et al., 2003). Ces modèles semblent offrir un cadre théorique intéressant pour étudier l'adoption d'un produit technologique. De plus, un questionnaire est disponible pour ces deux modèles.

4. Les déterminants affectivo-motivationnels

Un certain nombre de variables rattachées à la sphère affectivo-motivationnelle a été intégré dans les différents modèles de l'acceptabilité précédemment présentés afin de saisir plus largement le jugement des utilisateurs envers un système. Cependant, il semble pertinent de consacrer une place particulière au rôle de ces variables et d'en analyser plus profondément leurs spécificités. Ainsi, les variables les plus récurrentes dans la littérature seront présentées dans les parties suivantes.

L'ajout de variables affectivo-motivationnelles aux modèles de l'acceptabilité vise à améliorer leurs pouvoirs explicatifs et prédictifs. Ces dimensions visent à capturer des éléments du jugement utilisateur qui vont au-delà de l'évaluation rationnelle des qualités du système. Ainsi, l'effet des émotions sur le jugement a fait l'objet de nombreux travaux.

4.1. Les émotions

Il est aujourd'hui bien reconnu que les émotions peuvent jouer un rôle important dans des domaines variés du fonctionnement humain (e.g., Pessoa, 2008).

4.1.1. Cadre théorique

De nombreuses définitions de l'émotion ont été proposées au sein de la littérature qui peut notamment s'expliquer par la variété des théories dans le domaine. Kleinginna et Kleinginna (1981) ont par exemple recensé une liste de 92 définitions, témoignant de la diversité des caractérisations des émotions. Les définitions proposées par Keltner et Gross (1999, p. 468) semblent pertinentes pour les présents travaux : « *pattern biologique relativement court et épisodique de perceptions, d'expériences, de réactions physiologiques, d'actions et de communications qui répondent à des événements physiques et sociaux spécifiques* » [notre traduction][4]. Au-delà de la définition du concept d'émotion, il existe plusieurs approches des émotions (se référer au *handbook* édité par Lewis, Haviland-Jones et Barrett (2010)

[4] La définition proposée par Scherer (2005, p. 697) apparaît également comme pertinente : « *an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism* ».

pour une revue plus complète). Seule l'approche basée sur l'évaluation cognitive (*appraisal)* rattachée au modèle dit « des processus composants » sera exposée. En effet, cette approche semble, à l'heure actuelle, la plus commune (Moors, Ellsworth, Scherer, & Frijda, 2013).

Selon l'approche des processus composants, les émotions sont considérées comme un processus à cinq composantes : le sentiment subjectif, les changements psychophysiologiques, l'expression motrice, les tendances à l'action et l'évaluation cognitive. Il est couramment admis que l'évaluation cognitive est la dimension qui détermine les changements dans les quatre autres composantes (Sander & Scherer, 2014). Enfin, selon Scherer (2005), les émotions auraient sept caractéristiques spécifiques :

- Elles apparaitraient à la suite d'un stimulus ;
- Il y aurait une synchronisation des cinq composantes présentées précédemment ;
- Elles influenceraient le comportement ;
- Elles auraient une évolution rapide ;
- Des variations dans leurs intensités seraient possibles ;
- Elles seraient de durée relativement courte ;
- Elles seraient guidées par l'évaluation cognitive.

Parmi ces éléments, l'évaluation cognitive aurait un rôle essentiel dans la génération des émotions. Elle représente un processus cognitif d'évaluation d'une situation par un individu au regard de ses capacités (Jokinen, 2015) et renvoie, par conséquent, à la signification qu'une personne donne à un événement (Kaiser, Wehrle, & Schenkel, 2013).

4.1.2. Les émotions en IHM

Depuis plusieurs décennies, les chercheurs s'intéressent à l'influence des émotions sur le comportement humain et ont démontré des effets importants dans des domaines très variés. On retrouve par exemple une influence sur la satisfaction au travail (e.g., Weiss, Nicholas, & Daus, 1999), sur la prise de décision (e.g., Lerner, Li, Valdesolo, & Kassam, 2015), etc. L'intérêt pour les émotions s'est étendu aux

Interactions Homme-Machine (IHM) où le nombre de recherches portant sur le lien entre émotions et IHM ne cesse d'augmenter (Lim et al., 2008). En effet, les émotions, reflets de nos perceptions, sont devenues une dimension centrale de l'expérience utilisateur (Desmet, 2002).

Plusieurs approches ont été employées pour étudier le rôle des émotions dans le cadre des interactions avec un système. Cependant, ces approches ont conduit à la cohabitation des termes émotions, humeurs et affects sans toujours de distinction sur le plan conceptuel notamment dans le domaine des IHM. Or, ces concepts ne sont pas pour autant synonymes (Luminet, 2008). En effet, l'humeur (e.g., gai, déprimé) est généralement considérée comme un état affectif caractérisé par la manifestation relativement longue (de plusieurs heures à plusieurs jours) de certains types de ressentis subjectifs qui influencent notamment le comportement de l'individu (Scherer, 2005). La vitesse de déclenchement de l'humeur est également plus longue que pour les émotions (Thayer, 2010) avec généralement une faible intensité. De plus, l'humeur peut apparaître sans cause apparente qui peut être clairement reliée à un événement.

Une première partie de la littérature porte ainsi sur l'impact de l'humeur préalable à l'interaction sur l'acceptabilité. Par exemple, une humeur positive pourrait augmenter l'acceptabilité d'un produit (Djamasbi & Strong, 2008). Venkatesh et Speier (1999) ont mis en avant l'influence significative de l'humeur sur la motivation intrinsèque et l'intention d'usage. Bien que cette approche soit intéressante, l'humeur est un élément lié à l'individu et n'est pas forcément influencée par les propriétés du système. Or, ce sont les facteurs sur lesquels il est possible d'influer qui présentent le maximum d'intérêt dans le cadre des travaux de cette thèse. Ainsi, seuls les travaux en lien avec les émotions (i.e., réactions multi-componentielles à un événement) semblent pertinents.

Une seconde approche vise à étudier l'influence des caractéristiques d'un système sur les émotions ressenties par les utilisateurs au cours de l'interaction ainsi que sur l'adoption d'un produit. Cette approche étant reliée à la perception des qualités des systèmes par les utilisateurs, elle semble la plus pertinente au regard des travaux de cette thèse. Cependant, peu de recherches ont actuellement été conduites. De plus, il n'existerait toujours pas de socle théorique adapté à la prise en compte des émotions dans les IHM (Saariluomaand & Jokinen, 2014). Néanmoins des résultats intéressants sont à noter dans la littérature. Pour commencer, Mahlke et Minge (2008)

ont tenté d'appliquer le modèle des processus composants de Scherer (2005) aux IHM dans l'optique de mettre en lumière les dimensions indispensables pour comprendre les réactions émotionnelles des utilisateurs. À partir de leurs résultats, ils concluent que la prise en compte de plusieurs composantes peut améliorer la finesse d'analyse des émotions mais expose au risque d'obtenir des résultats contradictoires. Ainsi, il n'y aurait pas d'approche universelle pour les travaux sur les émotions en IHM et la sélection des facteurs pertinents doit se faire au cas par cas. Sur la même base théorique des émotions, Éthier, Hadaya, Talbot, et Cadieux (2008) ont mis en avant l'influence des caractéristiques d'une interface sur le jugement cognitif (*appraisal*) de l'utilisateur. D'autres études se sont intéressées à l'impact de l'esthétisme et de l'utilisabilité (« degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié » ; NF EN ISO 9241-11, 1998, p. 2) sur les réactions émotionnelles. Ainsi, Porat & Tractinsky (2012) ont montré l'influence positive de l'esthétisme sur les dimensions plaisir et activation (*arousal*) dans le cadre de site web de e-commerce. Mahlke & Lindgaard (2007) ont, quant à eux, démontré un effet de l'utilisabilité sur les réactions émotionnelles (sentiments subjectifs et évaluations cognitives).

Si l'on s'intéresse à d'autres domaines que les produits technologiques, comme, par exemple, celui du comportement client, la prise en compte des affects passe souvent à travers la notion de satisfaction (e.g., Cronin Jr., Brady, & Hult, 2000 ; White & u, 2005). Cette approche se retrouve également dans le domaine des IHM (e.g., Lindgaard & Dudek, 2003 ; Tractinsky & Zmiri, 2008). Hassenzahl (2004, 2005a) considère notamment la satisfaction comme une conséquence émotionnelle de l'usage d'une interface. Les émotions sont également parfois considérées comme un élément de la satisfaction dans le cadre des technologies (De Guinea & Markus, 2009). Cependant, l'utilisation du concept de satisfaction pose problème pour deux principales raisons. Tout d'abord, cela revient à considérer que le jugement de l'utilisateur se fait uniquement sur les qualités instrumentales d'un produit (Février, Gauducheau, Jamet, Rouxel, & Salembier, 2011). Or, cette hypothèse ne semble pas réaliste (Bargas-Avila & Hornbaek, 2011). De plus, la satisfaction apparaît comme un concept aux contours flous parfois considéré comme la résultante d'autres dimensions

(e.g., Calisir & Calisir, 2004) et d'autres fois comme une composante de l'intention comportementale (e.g., Limayem, Hirt, & Cheung, 2007).

En conclusion, bien qu'il y ait un réel intérêt pour la prise en compte des affects dans les IHM, les recherches ayant exploré l'influence de ces caractéristiques sur les réactions émotionnelles restent encore limitées. Cela pourrait notamment s'expliquer par l'absence de cadre théorique adapté à ces questions (Gauducheau, 2009 ; Saariluomaand & Jokinen, 2014).

4.1.3. Outils de mesure des émotions

En ce qui concerne l'évaluation des émotions, différents outils de mesure sont présents dans la littérature, notamment en fonction des cadres théoriques.

Dans le cadre de l'approche dimensionnelle, l'outil Self-Assessment Manikin (SAM ; Bradley & Lang, 1994) permet notamment d'évaluer la Valence, l'Arousal et la Dominance (voir pour un exemple d'échelle), l'Affect Grid (Boyle, 1984), les Emocards (Desmet, Overbeeke, & Tax, 2001).

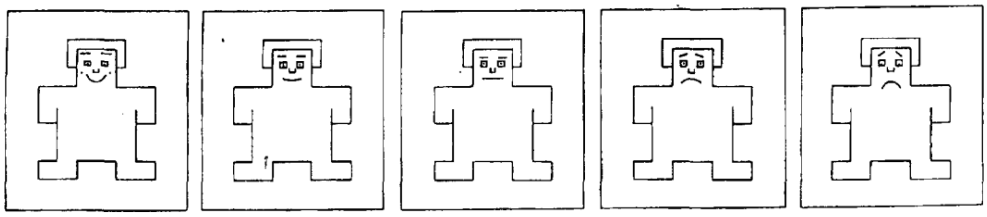


Figure 6. Echelle de valence extraites du SAM (Bradley & Lang, 1994)

Dans l'approche catégorielle, on retrouve plusieurs outils comme la PANAS (Watson, Clark, & Tellegen, 1988), la Differential Emotion Scale (Watson, Clark, & Tellegen, 1988) et la Computer Emotion Scale (Kay & Loverock, 2008).

Conjointement à la prise en compte des réactions affectives, des facteurs motivationnels comme le plaisir perçu auraient une influence sur l'intention d'usage envers un produit technologique.

4.2. Motivation, plaisir perçu et fun

Afin d'améliorer le modèle du TAM, plusieurs auteurs (e.g., F. D. Davis et al., 1992) ont entrepris des recherches sur les variables à intégrer pour en améliorer le

pouvoir explicatif. Une partie de ces recherches a été articulée autour de la notion de motivation, de qualité hédonique ou encore de plaisir perçu.

4.2.1. Motivation

Pour commencer, le champ théorique de la motivation est vaste, et par conséquent seul le cadre théorique proposé par Ryan et Deci (2000) sera présentement exposé. Celui-ci propose un découpage entre la motivation intrinsèque et la motivation extrinsèque (R. M. Ryan & Deci, 2000). La motivation intrinsèque renvoie au fait de réaliser une action parce qu'elle est intéressante en soi. L'action n'est pas réalisée dans l'espoir d'obtenir quelque chose en retour mais c'est le fait de réaliser l'action qui est une motivation en soi. La motivation extrinsèque renvoie, quant à elle, au fait de faire quelque chose pour une raison extérieure.

La motivation, matérialisée par la mesure de ses composantes intrinsèque et extrinsèque (F. D. Davis et al., 1992), est régulièrement employée dans le domaine des IHM notamment dans les travaux sur l'acceptabilité. En effet, ces composantes servent de variables explicatives à l'intention d'usage (e.g., Venkatesh, Speier, & Morris, 2002). Néamoins, une large partie des travaux utilisant les concepts de motivation reposent sur le concept de plaisir perçu.

4.2.2. Plaisir perçu (perceived enjoyment)

Davis et ses collaborateurs (1992) ont défini le concept de plaisir perçu (*perceived enjoyment*) comme étant une forme de motivation intrinsèque à l'usage d'un ordinateur. Ainsi, il est ainsi défini comme étant le degré selon lequel l'utilisation d'un système est considérée comme agréable en dehors de la performance du système. La notion de plaisir perçu est souvent associée à celle de facilité d'usage. De fait, le plaisir perçu est soit considéré comme un antécédent soit comme une résultante de la facilité d'usage perçue (Sun & Zhang, 2006). Certains auteurs ont montré une influence du plaisir perçu sur l'intention de retourner sur un site e-commerce (Koufaris, 2002) ou encore sur un site mobile (Cyr, Head, & Ivanov, 2006).

Enfin, cette variable revêt des appellations différentes en fonction des auteurs et des champs théoriques. Ainsi, ce concept peut se décliner en motivation hédonique (e.g., Venkatesh et al., 2012a), motivation intrinsèque (e.g., F. D. Davis et al., 1992) ou encore fun (e.g., Bruner & Kumar, 2005). Malgré cette diversité d'appellations, les items utilisés pour mesurer ces variables sont généralement ressemblants (e.g., Childers, Carr, Peck, & Carson, 2001 ; Cyr et al., 2006 ; F. D. Davis et al., 1992 ; S.

Davis & Wiedenbeck, 2001 ; van Schaik & Ling, 2011) : par exemple, « je trouve qu'utiliser [ce système] est agréable » [notre traduction] (F. D. Davis et al., 1992). Bien que certaines variantes apparaissent dans la définition, ces variables peuvent être considérées comme similaires.

Ainsi, le plaisir perçu (*perceived enjoyment*) est régulièrement considéré comme un indicateur de la motivation intrinsèque (Lee, Cheung, & Chen, 2005 ; Sun & Zhang, 2008). D'autres auteurs cherchent à prendre en compte la notion de motivation extrinsèque à travers la facilité d'usage perçue et/ou l'utilité perçue (e.g., Teo, Lim, & Lai, 1999 ; Venkatesh et al., 2012a).

Dans le prolongement des travaux liés à la motivation, des recherches ont conduit au développement d'un champ de recherche autour du concept d'engagement.

4.3. Engagement

La notion d'engagement n'est pas propre aux IHM. Elle se retrouve, par exemple, dans le domaine de l'éducation (e.g., Klem & Connell, 2004) ou du travail (e.g., Bakker, 2008a ; Schaufeli, Salanova, González-romá, & Bakker, 2002).

L'engagement peut se définir comme étant *« la qualité de l'expérience utilisateur qui met l'accent sur les aspects positifs de l'interaction, en particulier le fait d'être captivé par l'application web ainsi qu'être motivé par son usage* » [notre traduction] (Lehmann, Lalmas, om-Tov, & Dupret, 2012, p. 164). Cette variable jouerait un rôle important (Attfield et al., 2011) dans la Conception Centrée Utilisateur (CCU - i.e., « approche de conception et de développement de systèmes ayant pour objectif d'améliorer l'utilisabilité des systèmes interactifs en se concentrant sur l'utilisation du système concerné, et en appliquant les connaissances et techniques existantes en matière de facteurs humains/d'ergonomie et d'utilisabilité » ; NF EN ISO

9241-210, 2010)[5] et l'expérience utilisateur (UX - i.e., « perceptions et réactions d'une personne qui résultent de l'utilisation effective et/ou anticipée d'un produit, système ou service » ; NF EN ISO 9241-210, 2010). Cependant, à l'instar de l'UX, l'engagement reste un concept avec une définition évasive (Rozendaal, Keyson, & Ridder, 2009). Selon les auteurs, cela pourrait renvoyer à un état cognitif, un état affectif ou encore un comportement (O'Brien & Toms, 2010). De nombreux chercheurs ont tenté d'établir une liste des éléments impliqués dans l'engagement. L'attention focalisée (Attfield et al., 2011), la motivation (O'Brien & Toms, 2010) ou encore l'intérêt (Webster & Ho, 1997) seraient ainsi reliés à l'engagement. Cependant, peu de modèles généralistes ont été proposés dans la littérature. Celui d'O'Brien et Toms (2008), appliqué au IHM, semble néanmoins le plus connu et décompose le processus d'engagement en quatre étapes (voir figure 7) :

- **Le point d'engagement** : le processus d'engagement débute en réponse à l'interaction avec un système. Les composantes de cette interaction vont conduire l'utilisateur à s'engager avec le système. Les composantes suivantes seraient rattachées à cette étape : esthétisme, nouveauté, intérêt, motivation et buts.
- **La période d'engagement** : au cours de cette période, l'attention et l'intérêt de l'utilisateur sont maintenus à un niveau élevé grâce à un sentiment de contrôle, une perception de nouveauté et de challenge. Les composantes suivantes seraient reliées à cette étape : esthétisme, attention, conscience, contrôle,

[5] Selon la norme ISO 9241-210 (2010), la CCU repose sur 6 principes : 1) la conception est fondée sur une compréhension explicite des utilisateurs, des tâches et des environnements ; 2) les utilisateurs sont impliqués dans la conception et le développement ; 3) la conception est dirigée et précisée par l'évaluation centrée sur l'utilisateur ; 4) le processus est itératif jusqu'à obtention du résultat souhaité ; 5) la conception couvre l'expérience de l'utilisateur dans son intégralité, c'est-à-dire des aspects liés à l'efficacité et la performance aux aspects liés par exemple à la satisfaction et aux attentes des utilisateurs ; 6) l'équipe de conception inclut des compétences et des points de vue pluridisciplinaires.

interactivité, nouveauté, challenge, feedback, intérêt, affect positif et temps passé avec le système.

- **Le désengagement** : correspond au moment où l'utilisateur interrompt son utilisation de sa propre initiative ou pour des raisons externes. Ces composantes sont liées à : l'utilisabilité, le challenge, les affects positifs et négatifs et l'interruption.
- **Le réengagement** : parfois à la suite d'un désengagement, les utilisateurs recommencent le processus d'engagement.

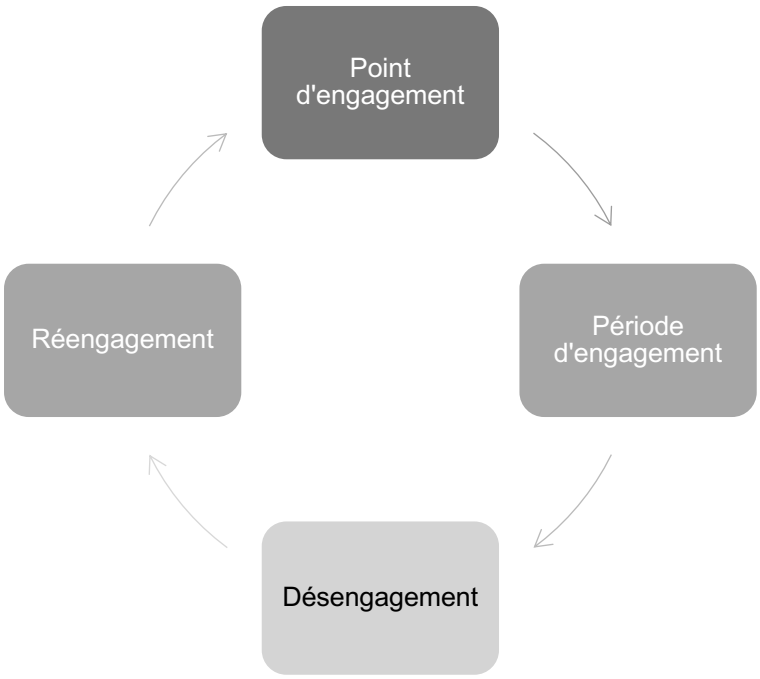


Figure 7. Modèle d'O'Brien & Toms (2008)

Par la suite, O'Brien et Toms (2010) ont tenté de valider leur modèle à partir d'une modélisation en équations structurelles. Ils ont ainsi retenu le modèle suivant : voir figure 8.

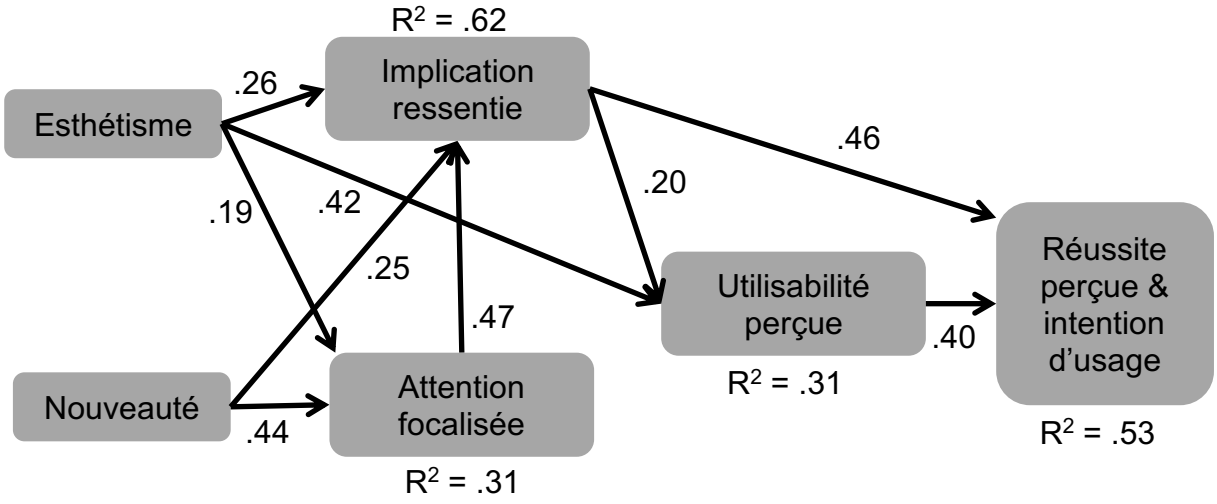


Figure 8. Modèle de l'engagement selon O'Brien & Toms (2010)[6]

Ainsi, de nombreuses variables de l'engagement semblent jouer un rôle dans la construction de l'avis de l'utilisateur. Cette influence est également relevée dans d'autres recherches : par exemple l'engagement aurait un rôle dans la préférence pour un site web (Hartmann, Sutcliffe, & Angeli, 2008) ou encore sur l'auto-efficacité perçue en informatique, qui est, elle-même, un antécédent aux variables du TAM (Scott & Walczak, 2009). Simultanément, l'intérêt pour le *flow*, concept partiellement en lien avec la notion d'engagement, s'est développé dans la littérature.

4.4. Le *Flow*

4.4.1. L'engagement ultime ?

Un autre concept, le *flow*, a pris de l'ampleur au cours des dernières années. Ce concept, que l'on doit à Csikszentmihalyi (2014), décrit une expérience optimale (équilibre entre les objectifs et les compétences) caractérisée par un engagement complet au sein d'une activité avec un intérêt. Pour définir le *flow*, Nakamura & Csikszentmihalyi (2002) se sont intéressés aux caractéristiques inhérentes à l'apparition de la sensation de *flow*. Selon eux, elles seraient au nombre de huit :

[6] R^2 : Variance expliquée. Ce chiffre, entre 0 et 1, permet de déterminer à quel point les prédicteurs permettent d'expliquer la variable à prédire. Par exemple, dans ce modèle, la variance de la variable « attention focalisée » est expliquée à 31 % (R^2 = .31) par « la nouveauté » et « l'esthétisme ».

- La perception d'un challenge correspondant aux compétences de l'individu ;
- La présence d'objectifs précis ;
- La présence de feedbacks immédiats ;
- Une concentration intense ;
- Une perte de la conscience de soi ;
- Une sensation de contrôle ;
- Une distorsion de la perception du temps ;
- L'activité est source de satisfaction en soi.

Cependant, seulement trois éléments sont généralement retenus dans les recherches en lien avec le *flow* (Bakker, 2008a) : l'absorption (i.e., immersion), le plaisir (*enjoyment*) et la motivation intrinsèque. Ainsi, de par sa nature, le *flow* est considéré comme une composante de la motivation intrinsèque (Agarwal & Karahanna, 2000). Dans le domaine des technologies, le *flow* est parfois considéré comme une forme d'engagement (Chapman, Selvarajah, & Webster, 1999 ; Webster & Ho, 1997). Le *flow* serait alors une forme d'engagement ultime avec un système.

Cet état particulier a été étudié dans des domaines variés comme le sport (e.g., Swann, Keegan, Piggott, & Crust, 2012) ou encore le travail (e.g., Eisenberger, Jones, Stinglhamber, Shanock, & Randall, 2005). En ce qui concerne le domaine des IHM, le *flow* est notamment utilisé pour améliorer le pouvoir explicatif du TAM (e.g., I. Ha, oon, & Choi, 2007). Ainsi, Davis & Wiedenbeck (2001) ont montré une influence positive (β = .58[7]) du *flow* sur la facilité d'usage perçue et l'utilité perçue. Hsu et Lu (2004) ont démontré dans le domaine du jeu vidéo un effet positif modéré (β = .12[7]) de l'expérience de *flow* sur l'intention d'usage.

4.4.2. Outil de mesure du flow

Pour évaluer le *flow*, plusieurs échelles ont été développées dans la littérature. On retrouve notamment le *WOrk-reLated Flow inventory* (WOLF) principalement appliqué au bien-être au travail (3 facteurs et 13 items ; Bakker, 2008b), la *Flow*

[7] Coefficients bêta standardisés.

experience (3 items ; Hsu & Lu, 2004), l'*Intensity of Flow* (2 dimensions et 9 items ; S. Davis & Wiedenbeck, 2001) ou la *Flow State Scale* (9 dimensions et 36 items - FSS ; S. A. Jackson & Marsh, 1996) / FFS-2 (9 dimensions et 36 items ; S. A. Jackson & Eklund, 2002).

Enfin, en parallèle des variables affectivo-motivationnelles présentées précédemment, les facteurs confiance et crédibilité perçue semblent avoir une importance dans l'intention d'usage envers un système.

4.5. Confiance, fiabilité et crédibilité

4.5.1. Définitions

Pour commencer, la crédibilité est considérée comme une qualité perçue par l'utilisateur. Elle peut prendre plusieurs formes différentes (Fogg & Tseng, 1999) :

- La crédibilité présumée qui repose sur des suppositions générales de l'utilisateur ;
- La crédibilité de réputation basée sur les éléments transmis par d'autres personnes sur le produit ;
- La crédibilité de surface qui prend appui sur une simple inspection du système ;
- La crédibilité liée à l'expérience avec le produit.

Ainsi, Fogg et ses collaborateurs (2003) se sont intéressés aux éléments à partir desquels les personnes construiraient leur avis sur la crédibilité d'un site web. Sur un total d'environ 2500 participants, ces derniers ont déclaré se baser en premier sur le visuel du produit (*design look*) puis sur la structuration de l'information.

En ce qui concerne la confiance (*trust*), elle peut être définie comme étant « *la croyance positive que l'on a, à propos de la fiabilité perçue d'un système, d'une personne ou d'un processus* » [notre traduction] (Fogg & Tseng, 1999, p. 81). Dans le cadre de l'acceptabilité de sites de e-commerce, la confiance semble ainsi avoir une influence positive sur les variables du TAM et l'intention d'achat (Jarvenpaa, Tractinsky, & Vitale, 2000 ; Pavlou, 2003). Ce résultat a été confirmé à plusieurs reprises (e.g., Corbitt, Thanasankit, & i, 2003 ; Gefen, Karahanna, & Straub, 2003).

4.5.2. Outil de mesures de la confiance

Plusieurs échelles ont été proposées pour évaluer la confiance perçue envers un produit technologique, on retrouve par exemple l'échelle de Gefen, Karahanna, &

Straub (5 items ; 2003) ou encore celle de Corbitt, Thanasankit & Hit (3 dimensions et 10 items ; 2003).

4.6. Superposition des variables affectivo-motivationnelles

Au sein de ce chapitre, un aperçu des travaux autour des facteurs affectivo-motivationnels en IHM a été proposé. Ces recherches font notamment appel à plusieurs champs théoriques de la psychologie (e.g., émotions, motivation). Cependant, certains concepts semblent se chevaucher sur le plan conceptuel au point d'être parfois mesurés de manière identique. Ainsi, les termes affect, humeur et émotions sont, par exemple, régulièrement utilisés sans distinction. Or, ces concepts présentent des différences importantes de définitions et doivent donc être distingués. A l'inverse, le concept de motivation intrinsèque est utilisé à travers la variable plaisir perçu. Néanmoins, les outils de mesure reprennent des items de motivation intrinsèque. Ainsi, le concept de motivation intrinsèque sera utilisé plutôt que le concept de plaisir perçu. En résumé, il a été tenté au sein de ce document de cloisonner les variables au mieux en s'appuyant sur les concepts théoriques princeps.

En Résumé

Au-delà des variables rattachées aux modèles d'acceptabilité TAM et UTAUT, les facteurs affectivo-motivationnels semblent être déterminants pour mieux appréhender les raisons de l'adoption d'un produit technologique. Cependant, ces variables sont encore rarement prises en compte notamment de par la complexité de leur évaluation et des cadres théoriques qui y sont liés.

5. Modèles du jugement utilisateur et de l'expérience utilisateur (UX)

Au-delà des modèles d'acceptabilité/d'acceptation et des facteurs affectivo-motivationnels ayant été ajoutés au cours du temps, d'autres recherches ont été conduites sur des composantes plus larges du jugement utilisateur. Deux approches différentes sont ainsi présentes dans la littérature : la première approche, la plus ancienne, s'appuie sur la notion d'utilisabilité du produit et ne sera que très rapidement exposée. La deuxième repose sur l'expérience utilisateur (UX), un concept plus récent visant à dépasser la notion d'utilisabilité.

5.1. La notion d'Utilisabilité

5.1.1. Définition

La notion d'utilisabilité a fait son apparition dans les années 80 (James R. Lewis, 2012) et a été définie par la norme ISO 9241-11 (1998, p. 2) comme étant le « *degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié* ». Cette notion a été au cœur des processus de conception pendant de nombreuses années (Kuutti, 2010) et doit son essor au développement des produits technologiques. Les travaux autour de la notion d'utilisabilité au sein de la norme ISO sont centrés sur les qualités fonctionnelles du produit. Bien que cette norme prenne en compte la dimension de satisfaction qui inclut, par exemple, la notion d'affect (Nielsen, 1994), elle est majoritairement mesurée par des questionnaires interrogeant les qualités fonctionnelles du système (Hornbæk & Law, 2007).

Cependant, bien que la notion d'utilisabilité soit nécessaire pour une utilisation facilitée des systèmes, elle n'apparait plus suffisante pour développer des produits de qualité. En effet, aujourd'hui le niveau des qualités fonctionnelles des produits a suffisamment augmenté pour s'intéresser à des qualités plus larges (e.g., esthétisme) (Robert & Lesage, 2011a). Ainsi, il semble primordial d'utiliser cette variable comme base pour la construction de nouveaux modèles mesurant plus largement l'expérience des utilisateurs lors d'une interaction avec un système. Ceci est le point de départ du concept de l'UX, détaillé au sein de la partie suivante.

5.1.2. Outil de mesure de l'utilisabilité

L'utilisabilité semble être l'élément pour lequel le plus d'échelles ont été développées. On y retrouve des échelles courtes unidimensionnelles comme le SUS (10 items ; Brooke, 1996), des questionnaires plus complets comme le CSUQ (3 dimensions et 19 items ; J. R Lewis, 1995) ou encore des outils plus récents comme l'UMUX (4 items ; Finstad, 2010) ou la DEEP (6 dimensions et 19 items ; T. ang, Linder, & Bolchini, 2012). Cependant, les analyses et qualités psychométriques des échelles de la littérature font encore aujourd'hui débat (e.g., Bosley, 2013 ; Cairns, 2013 ; J. R. Lewis, 2013).

5.2. La notion d'expérience utilisateur (UX)

Le concept d'expérience utilisateur (UX - *User eXperience*) est né de la volonté d'aller au-delà des approches basées sur l'utilisabilité et les paradigmes reliés au travail et à la tâche (Hassenzahl & Tractinsky, 2006). Une définition de l'UX a été proposée en 2010 dans la norme ISO 9241-210 : « perceptions et réactions d'une personne qui résultent de l'utilisation effective et/ou anticipée d'un produit, système ou service ». Trois notes apportant des précisions sur les éléments influençant l'UX ont été ajoutées à cette définition :

- « L'expérience de l'utilisateur inclut toutes les émotions, convictions, préférences, perceptions, réactions physiques et psychologiques, comportements et réalisations de ce dernier, qui interviennent avant, pendant et après l'utilisation » ;
- « L'expérience de l'utilisateur est une conséquence de l'image de marque, la présentation, la fonctionnalité, les performances, le comportement interactif et les capacités d'assistance du système interactif ; de l'état intérieur et physique de l'utilisateur résultant d'expériences passées, de ses attitudes, de ses compétences et de sa personnalité ainsi que du contexte d'utilisation » ;
- « L'utilisabilité, lorsqu'elle est interprétée du point de vue des objectifs personnels des utilisateurs, peut comporter le type d'aspects perceptifs et émotionnels généralement associés à l'expérience de l'utilisateur. Les critères d'utilisabilité peuvent être utilisés pour évaluer les aspects de l'expérience de l'utilisateur ».

Cependant, tous les auteurs ne s'accordent pas sur cette définition (Hole & Williams, 2007 ; Lallemand, Gronier, et al., 2015). Cela a conduit à des modèles et des approches différents notamment à cause de la variété des disciplines qui s'y sont intéressées (Kuniavsky, 2009). Ainsi, l'UX a été majoritairement étudiée selon deux approches : l'approche holistique et l'approche multidimensionnelle. Les travaux de cette thèse visant notamment à la construction d'un outil de mesure multidimensionnel, seule une rapide introduction à l'approche holistique sera proposée. En effet, les auteurs, dans cette approche, supposent que l'expérience avec un produit est une entité indivisible (Wright, McCarthy, & Meekison, 2005), ce qui ne semble pas compatible avec l'approche sélectionnée dans cette thèse.

5.2.1. Approche holistique

Au sein de l'approche holistique, les chercheurs considèrent l'UX comme un tout indivisible et non reproductible (e.g., McCarthy & Wright, 2007). Selon ce courant, qui repose sur une vision phénoménologique de l'expérience, l'UX ne doit pas être envisagée comme un ensemble de variables qui viennent jouer un rôle dans la construction de l'expérience, mais comme un ressenti unique et propre à chaque individu (McCarthy & Wright, 2007). Seul le modèle de McCarthy et Wright sera présenté ici pour donner un aperçu des travaux dans cette approche. Ainsi, McCarthy et ses collaborateurs (McCarthy & Wright, 2007 ; Wright et al., 2005) ont proposé un modèle qui s'articule autour des interactions entre les éléments du modèle ainsi que la manière dont ces éléments s'enrichissent mutuellement. Pour cela, ils se basent sur la définition de l'expérience de John Dewey (2005) : un ensemble indivisible d'actes, de sensations, de pensées, de ressentis durant une certaine situation. Ainsi, leur modèle se décompose en deux parties : d'une part, les quatre éléments interdépendants qui composent l'expérience et d'autre part les processus qui permettent de « faire sens » (*make sense*), d'évaluer l'expérience (Kim, 2015). Ces quatre dimensions sont : la dimension sensorielle, la dimension compositionnelle, la dimension émotionnelle et la dimension spatio-temporelle. Plus précisément, les différentes dimensions sont définies de la manière suivante :

- La dimension **sensorielle** correspond à l'engagement sensoriel dans une situation. Elle est reliée au caractère palpable et viscéral de l'expérience.
- La dimension **compositionnelle** renvoie aux relations entre les différents éléments qui composent l'expérience.

- La dimension **émotionnelle** correspond aux ressentis que l'on gardera de l'expérience vécue. De plus, les émotions ne sont pas seulement le résultat d'une situation mais elle guide aussi nos actions.
- La dimension **spatio-temporelle** traduit le fait que toute expérience a lieu dans une situation spatio-temporelle particulière.

Au-delà de ces quatre dimensions, il existerait six processus permettant de « faire sens » à partir des éléments présentés précédemment (Wright, Wallace, & McCarthy, 2008) :

- L'anticipation : chaque expérience est influencée par son anticipation ;
- La connexion : première impression que l'on se fait sur un système ;
- L'interprétation : perceptions et analyses des possibilités offertes par le système et leurs conséquences ;
- La réflexion : évaluation de l'expérience en temps réel par l'individu ;
- L'appropriation : mise en lien de l'expérience vécue avec ses expériences précédentes ;
- Le rappel : lorsqu'un individu revit une expérience.

Enfin, l'expérience ressentie par chaque utilisateur est considérée comme évoluant au cours du temps et en fonction de ses autres expériences.

En conclusion, l'expérience lors de l'usage d'un produit est un ressenti unique et propre à chaque individu selon l'approche holistique. Ainsi, évaluer différentes dimensions perçues par les individus ne semble pas compatible avec cette théorie et par conséquent ne semble pas être la plus adaptée à cette thèse. Ainsi, la partie suivante sera consacrée à la présentation de l'approche multidimensionnelle de l'UX sur laquelle reposera cette thèse.

5.2.2. Approche multidimensionnelle

L'approche multidimensionnelle de l'UX, sans toutefois s'opposer à l'approche holistique, repose sur des fondements différents. Ici, les chercheurs tentent de distinguer les facteurs jouant un rôle dans l'UX (Hassenzahl, Diefenbach, & Göritz, 2010). Au sein de cette thèse, cette approche ancrée dans la méthodologie expérimentale (Hassenzahl & Monk, 2010) et différentielle offre, à notre sens, le

meilleur fondement pour construire un outil de mesure multidimensionnelle. Par conséquent, les principaux modèles de l'UX rattachés à cette approche seront présentés ci-dessous.

Modèle d'Hassenzahl

Selon Hassenzahl (2004), les individus percevraient les produits selon deux attributs : leurs qualités pragmatiques et leurs qualités hédoniques. Plus précisément, les qualités pragmatiques renverraient à la perception qu'ont les utilisateurs des capacités offertes par le produit pour atteindre le but recherché (« *do-goals* ») (Hassenzahl, 2008). Ces qualités sont axées sur le produit et sont à rapprocher des notions d'utilisabilité et d'utilité perçue. Quant aux qualités hédoniques, elles sont reliées à la perception que le produit permet d'atteindre des buts d'accomplissement (e.g., besoin d'être compétent) (« *be-goals* »). Ces attributs hédoniques seraient composés de trois éléments : identification, évocation et stimulation. Plus précisément, l'identification renverrait au besoin de s'exprimer soi-même à travers ses objets (Hassenzahl, 2004). Pour ce qui est de l'évocation, cela correspondrait à la capacité d'un produit à rappeler des souvenirs et des symboles[8]. Enfin, la stimulation peut être définie comme le besoin de challenge et de nouveauté chez les utilisateurs (Hassenzahl, 2004). En résumé, ces trois éléments reposeraient sur des besoins humains (Hassenzahl, 2005a) et sont ainsi centrés sur l'individu. La perception de ces qualités est influencée par plusieurs caractéristiques du système.

Ainsi, le contenu du système, sa présentation, ses fonctionnalités et les éléments d'interaction (voir figure 9) auraient une influence sur la perception des qualités perçues du système.

[8] La variable « évocation » (2005a) apparaît dans certaines versions du modèle d'Hassenzahl (e.g., Hassenzahl, 2005a). Cependant, Hassenzahl (2004) précise que cette variable n'est pertinente que lorsque le système « est en mesure de provoquer des souvenirs et agir comme un symbole du passé » [notre traduction].

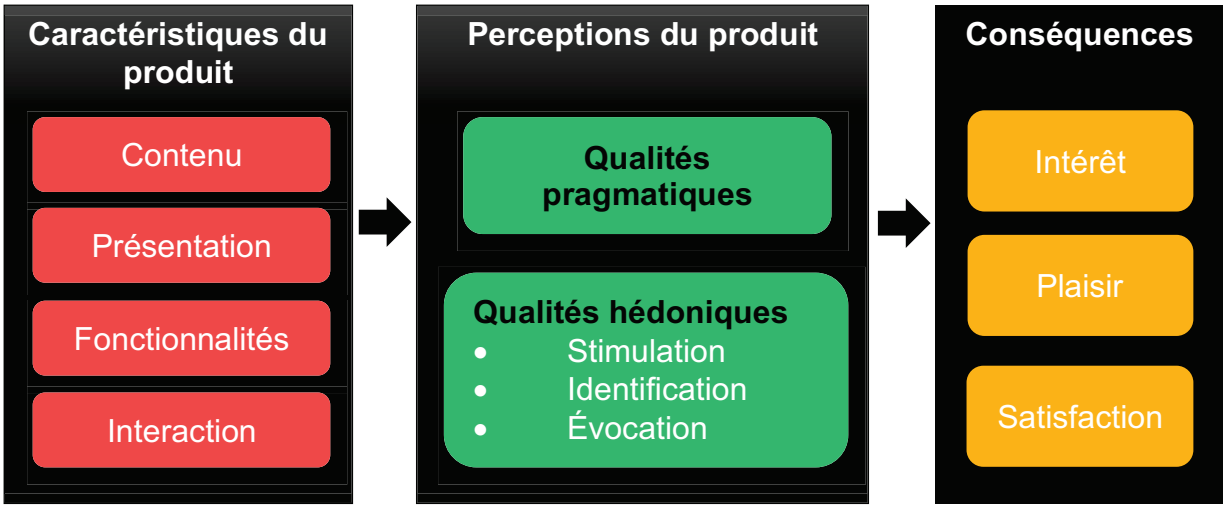


Figure 9. Modèle de l'UX selon Hassenzahl et Monk (2010)

Selon ce modèle, un élément supplémentaire, le mode d'usage, semble avoir un rôle dans l'évaluation du produit par l'utilisateur (Hassenzahl, Kekez, & Burmester, 2002). Il existerait ainsi deux modes d'usage : le mode but et le mode activité. Au sein du mode « but », les actions seraient guidées par un objectif à atteindre et le produit serait l'outil qui le permet (e.g., remplir sa déclaration d'impôts en ligne). Dans le cadre du mode « activité », l'utilisation du produit serait un but en soi (e.g., utiliser son nouveau téléphone pour en découvrir les fonctionnalités).

De par la présence d'un questionnaire associé, l'AttrackDiff (Hassenzahl, 2004), ce modèle est très régulièrement utilisé (Diefenbach, Kolb, & Hassenzahl, 2014) pour évaluer des produits (e.g., Obrist et al., 2012 ; Schaffer, Seebode, Wechsung, Metze, & Kühnel, 2009). Il est composé de 28 items répartis en 4 sous-dimensions (e.g., item pour les qualités pragmatiques : « je juge [ce système] comme : compliqué – simple » - voir Annexe 3 pour une présentation complète du questionnaire).

Cependant, ce modèle n'intègre pas explicitement les réactions émotionnelles de l'utilisateur et, à notre connaissance, n'a jamais fait l'objet d'analyses sur sa structure. Thüring et Mahlke (2007a) proposent, quant à eux, d'intégrer la dimension émotionnelle au sein des modèles UX.

Modèle de Thüring et Mahlke

Thüring et Mahlke (2007a) se sont intéressés à l'influence de l'utilisabilité perçue et de l'esthétisme sur le jugement d'interface au travers de trois études expérimentales. Au sein de celles-ci, les participants devaient évaluer des interfaces

de lecteurs audio ou de téléphones modifiées en fonction des conditions expérimentales (e.g., faible utilisabilité/ forte utilisabilité). La première étude visait à étudier les liens entre utilisabilité et émotions. Les participants devaient soit évaluer une interface mal conçue (faible utilisabilité) soit une interface bien conçue (forte utilisabilité). Les chercheurs ont constaté un effet de l'utilisabilité sur les émotions : l'interface bien conçue conduisait à des émotions plus positives et une activation (*arousal*) plus faible. Dans la seconde étude, le lien entre utilisabilité et esthétisme était exploré. Thüring et Mahlke ont noté un effet positif de l'esthétisme et de l'utilisabilité sur le jugement global avec une influence plus importante de l'utilisabilité. Enfin, la troisième étude était centrée sur les liens entre utilisabilité, esthétisme et émotions. Ils ont montré une influence positive de l'utilisabilité sur les qualités instrumentales et les émotions. De plus, il apparaît un effet positif de l'esthétisme sur les qualités non instrumentales et les émotions. À partir de ces résultats, Thüring et Mahlke (2007a) ont proposé un nouveau modèle de l'UX , le CUE-Model (voir figure 10).

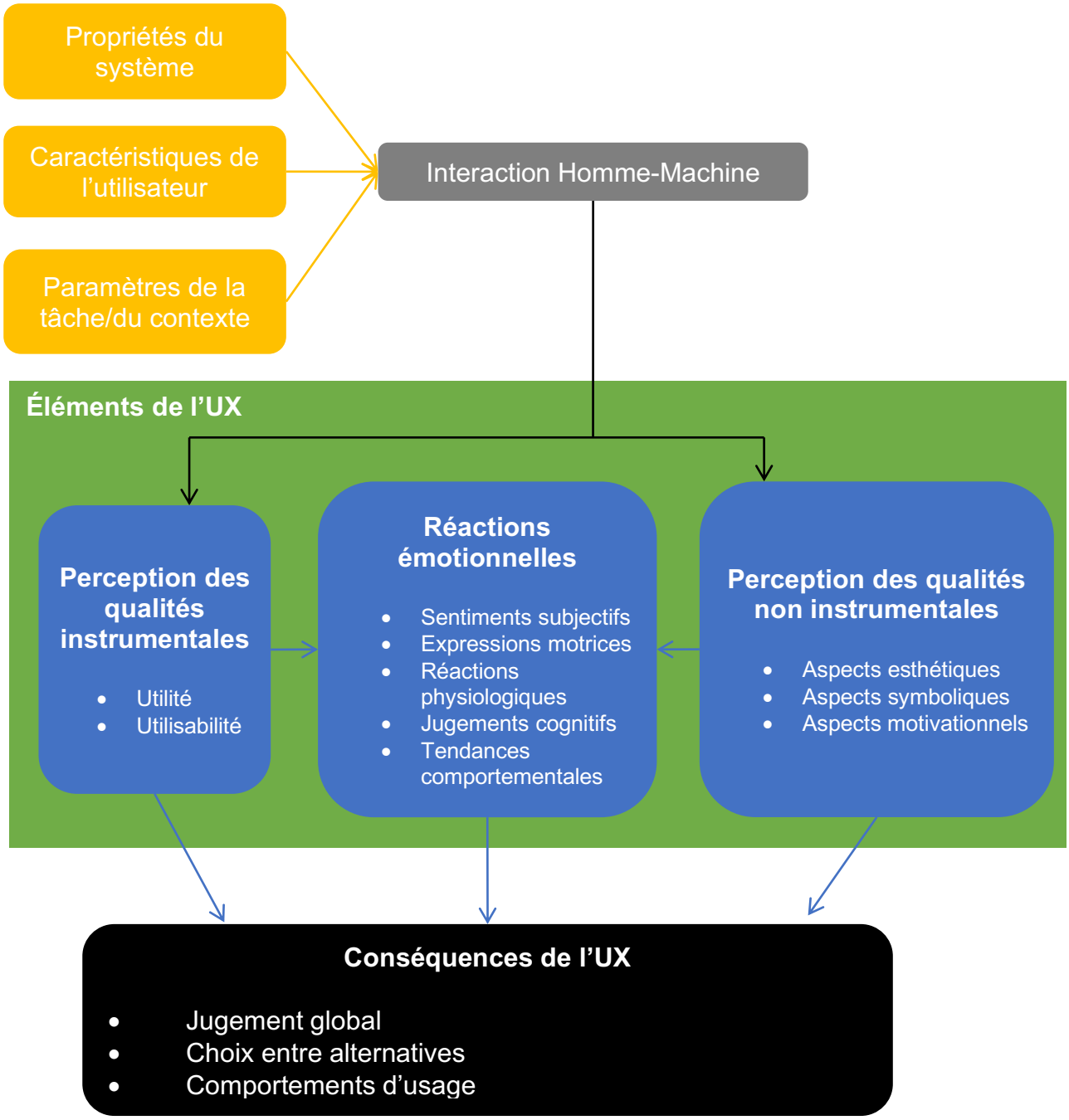


Figure 10. Modèle de l'UX selon Mahlke & Thüring (2007) : le CUE-Model

Selon ce modèle, l'Interaction Homme-Machine serait influencée par trois paramètres : les propriétés du système, les caractéristiques de l'utilisateur ainsi que le contexte. Cette interaction conditionne plusieurs éléments de l'UX. Pour commencer, elle aurait une influence directe sur la perception des qualités instrumentales et non instrumentales du produit. Ces deux éléments seraient à l'origine des réactions émotionnelles. Enfin, ces évaluations conditionneraient le jugement du produit et les comportements d'usage. On peut noter que, contrairement au modèle d'Hassenzahl

(voir Figure 9), les émotions sont considérées comme un élément de l'UX et non pas une conséquence de l'UX.

Ce modèle présente un intérêt lié à sa prise en compte des réactions émotionnelles ainsi que des qualités instrumentales et non instrumentales. Cependant, la structure de ce modèle reste théorique et n'a pas été testée, à notre connaissance, dans son ensemble. Néanmoins, un outil de mesure, le meCUE (Minge, 2013) a été proposé sur la base de ce modèle théorique. Il est composé de quatre dimensions (i.e., perception du produit, émotions, conséquences de l'usage et évaluation globale) pour un total 34 items. Les participants doivent se positionner sur des échelles de Likert en 7 points (e.g., « le produit est facile à utiliser » [notre traduction] pour évaluer la facilité d'usage).

Enfin, d'autres modèles, moins populaires, ont également été proposés dans la littérature, comme le modèle de Roto (2007) proposant un découpage de l'UX en trois éléments.

Autres modèles

D'autres modèles moins connus ont été proposés dans la littérature : le modèle de Roto (2007), le modèle de Arhippainen et Tähti (2003) ou encore celui de Robert et Lesage (Robert & Lesage, 2011b).

Premièrement, le modèle de Roto (2007) présuppose que l'expérience est propre à chaque individu et implique un produit ou un service avec lequel il est possible d'interagir. De plus, l'UX serait influencée à la fois par les attentes des utilisateurs ainsi que par des éléments extérieurs à l'interaction. Roto propose également une liste d'éléments qui rentrent en compte dans l'UX au cours de l'interaction (voir figure 11). Ainsi, l'UX serait le résultat de trois éléments : l'utilisateur, le système et le contexte. En premier lieu, l'utilisateur est composé de ses caractéristiques telles que ses attentes ou ses ressources. L'utilisateur va interagir avec le système, ce qui va conduire à une certaine perception du système. Cette interaction est influencée par le contexte tel que la tâche à réaliser.

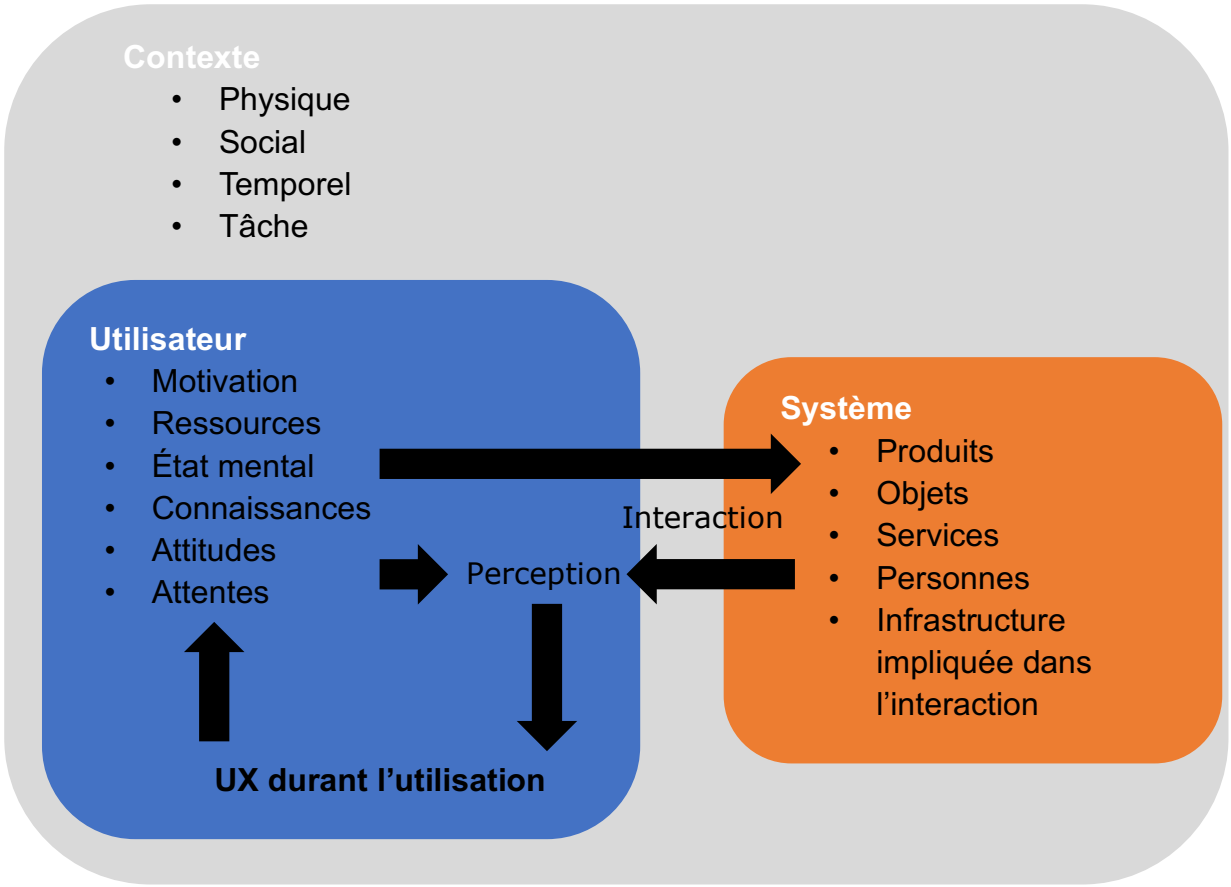


Figure 11. Éléments influençant l'UX au cours de l'interaction selon Roto (2007)

Arhippainen et Tähti (2003) considèrent, quant à eux, au sein de leur modèle que l'UX est influencée par cinq facteurs : l'utilisateur, les facteurs sociaux, les facteurs culturels, le contexte et le produit. Un ensemble de sous-variables est associé à chacun de ces facteurs (voir figure 12). À titre d'exemple, les valeurs, les émotions ou encore les attentes de l'individu sont reliées au facteur utilisateur.

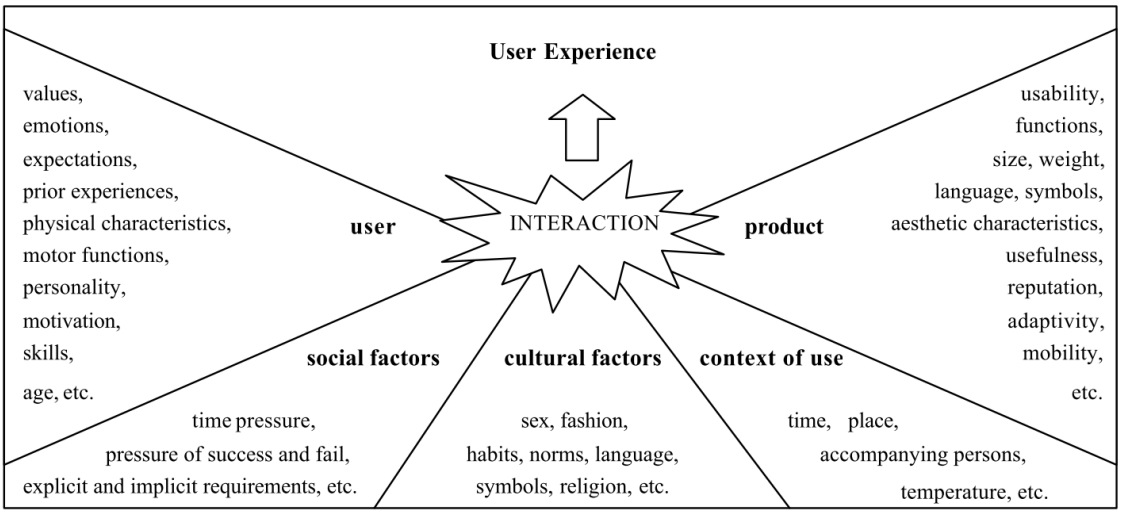


Figure 12. Éléments de l'UX selon Arhippainen et Tähti (2003)

Finalement, Robert et Lesage (Robert & Lesage, 2011b) ont proposé un modèle à mi-chemin entre l'approche holistique et multidimensionnelle. L'UX y est définie comme étant « *[…] un construit multidimensionnel qui définit l'effet global sur l'utilisateur de l'interaction avec un système dans un contexte particulier au cours du temps* » [notre traduction] (Robert & Lesage, 2011a, p. 323, 2011b). En résumé, selon ce modèle (voir figure 13), l'UX serait composée de six dimensions : fonctionnelle, physique, perceptuelle, cognitive, sociale, psychologique. De plus, l'UX est considérée comme subjective, dynamique, dépendante du contexte, à plusieurs niveaux de granularités et influencée par l'utilisateur, le système, l'activité ainsi que le contexte.

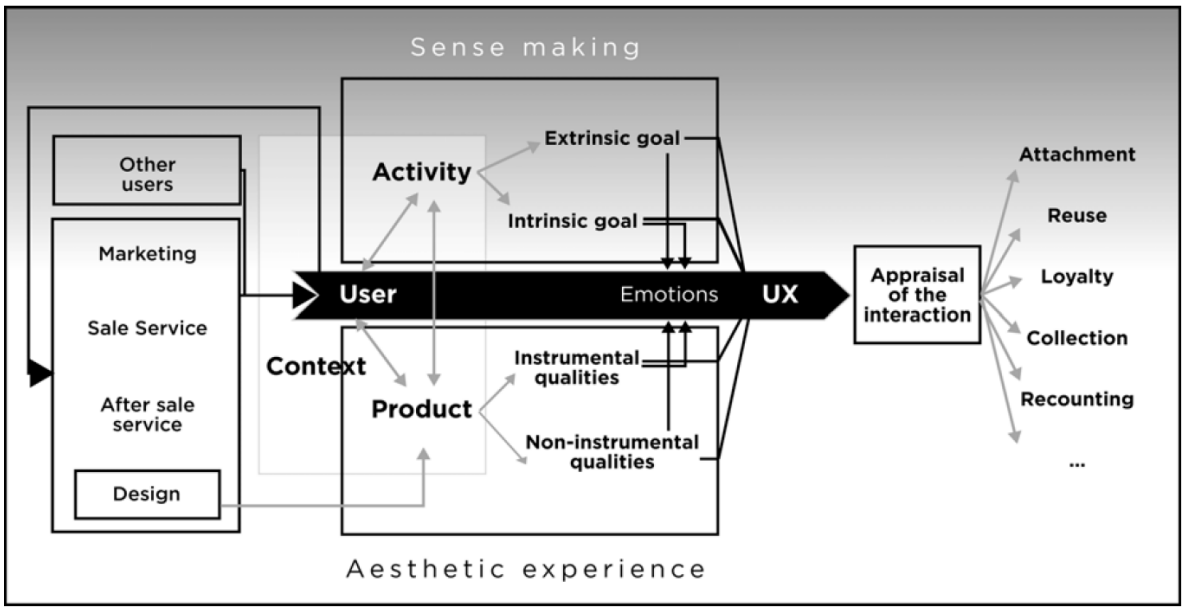


Figure 13. Modèle de Robert & Lesage (2011a)

En parallèle de ces modèles théoriques sur l'UX multidimensionnelle comme celui proposé par Robert et Lesage (2011a), le concept de qualité d'expérience s'est également répandu, notamment dans des domaines de recherches plus techniques.

5.3. Qualité d'expérience

Dans le domaine informatique, la notion de qualité d'expérience (*Quality of Experience* – QoE) a fait son apparition (voir Wechsung & De Moor (2014) pour une revue du concept de QoE). Ce concept prend lui aussi racine dans les travaux sur l'acceptabilité montrant que le bon fonctionnement d'un produit (*Quality of Service* - QoS) n'est pas suffisant pour qu'il soit adopté (Microsoft Corporation, 2007 ; W. Wu et al., 2009). Selon cette conception, le bon fonctionnement ne correspond pas à un système avec une utilisabilité de qualité mais à un système fiable sur le plan technique. À la vue de l'augmentation des capacités techniques des produits, des chercheurs en informatique ont cherché à obtenir un cadre théorique prenant en compte également l'utilisateur (Fiedler, Hossfeld, & Tran-Gia, 2010). Plusieurs définitions de la QoE ont alors été proposées. On trouve, par exemple, celle de Wu et ses collaborateurs (2009, p. 483) : « *la QoE est un construit multidimensionnel reposant sur les perceptions et les comportements d'un utilisateur représentant, subjectivement et objectivement, son état émotionnel, cognitif et ses réponses comportementales lorsqu'il utilise un système* » [notre traduction]. Dans un objectif de standardisation, l'International

Telecommunication Union (ITU) (notre traduction, 2007, p. 2) a proposé de définir la QoE comme « l'*acceptabilité globale d'une application ou d'un service, telle qu'elle est subjectivement perçue par l'utilisateur final* ». Il est également ajouté que « [la] *QoE inclut les effets du système de bout en bout* » et que *« l'acceptabilité globale peut être influencée par les attentes de l'utilisateur ainsi que le contexte* ».

À notre connaissance, il n'existe pas de modèle générique de la QoE adapté à n'importe quel type de produit. Cependant, Wu et ses collaborateurs (2009) ont proposé un modèle de la QoE (voir figure 14) axé sur les environnements multimédia interactifs distribués (DIME). Bien que ce modèle soit spécialisé pour un type de système, la majorité des éléments contenus dans celui-ci pourra être appliquée à d'autres domaines. Cependant, il est important de noter que la structure de ce modèle n'a pas fait l'objet d'études à notre connaissance.

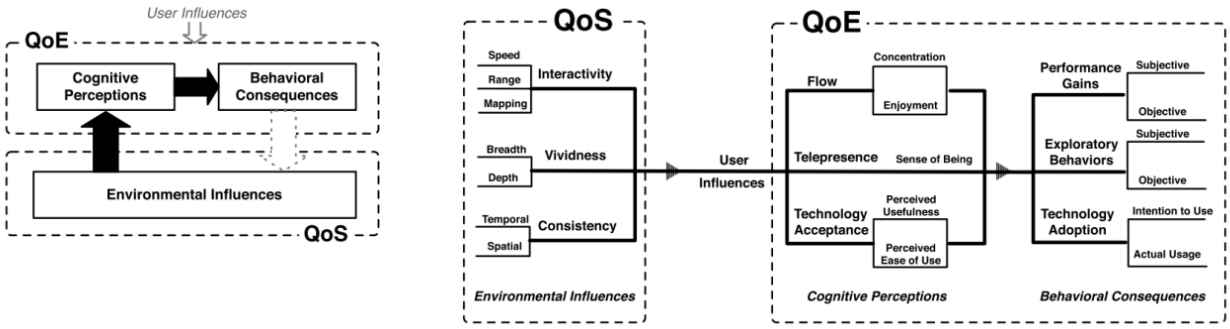


Figure 14. QoS et QoE pour les DIMEs (W. Wu et al., 2009)

Enfin, certains auteurs (Wechsung & De Moor, 2014) ont comparé la QoE et l'UX. Ces deux concepts sont ainsi considérés comme proches l'un de l'autre. Ils sont tous les deux issus de la nécessité d'aller au-delà des qualités techniques pour l'un et de l'utilisabilité pour l'autre. Ils ont également pour objectif d'intégrer l'avis de l'utilisateur au sein de l'évaluation. En revanche, ils se démarquent sur plusieurs points. Premièrement, ils proviennent de champs théoriques différents. L'UX provient du champ des IHM alors que la QoE provient du domaine des télécommunications. De manière plus importante, l'UX est subjective et individuelle ainsi que centrée sur l'utilisateur, alors que la QoE est largement dépendante de la qualité de service et centrée sur la technologie.

5.4. Outils de mesure de l'UX

Plusieurs types de mesures ont été développés afin d'évaluer l'UX (voir Lallemand & Gronier, 2015 pour une revue pratique des outils d'évaluation de l'UX). Deux questionnaires sont principalement utilisés dans la littérature. Tout d'abord, l'Attrackdiff qui repose sur le modèle théorique d'Hassenzahl (2004) a tout d'abord été proposé en allemand puis a été traduit en français par Lallemand et collaborateurs (2015). Ce questionnaire comporte 28 items répartis en 4 sous-dimensions (i.e., qualités pragmatiques, stimulation, identification et évaluation globale). Les participants se positionnent sur des échelles sémantiques différentielles en 7 points (e.g., la paire de mots « Original - Conventionnel » est proposée pour évaluer la stimulation). Les qualités psychométriques de l'échelle originale font l'objet de critiques (e.g., Lallemand, Koenig, et al., 2015). Même si la validation en français proposée par Lallemand, Koeniga, Gronier et Martin (2015) apporte des améliorations, les méthodes employées pourraient être améliorer : l'utilisation d'analyses factorielles confirmatoires, en plus des analyses factorielles en composantes principales déjà réalisées pourrait compléter les preuves sur les qualités psychométriques de l'échelle (J. R. Lewis, 2013). De plus, de par son rattachement au modèle d'Hassenzahl (2004), l'Attrackdiff inclut qu'un nombre réduit de variables rattachées à l'expérience utilisateur.

En parallèle, le meCUE (Minge, 2013), reposant sur le modèle de Thüring & Mahlke (2007b), a été construit. Il est également composé de quatre éléments (i.e., perception du produit, émotions, conséquences de l'usage et évaluation globale) avec un total 34 items. Les participants se positionnent sur des échelles de Likert en 7 points (e.g., « le produit est facile à utiliser » [notre traduction] pour évaluer la facilité d'usage). Les preuves apportées sur les qualités psychométriques du meCUE (Kuhr, 2013 ; Minge, Thüring, & Wagner, 2016), notamment sur la structure factorielle, de cette échelle, sont encore à confirmer par l'usage d'outils statistiques adaptés (e.g., analyses factorielles exploratoires et confirmatoires).

Un autre questionnaire est disponible dans la littérature, le UEQ (Laugwitz, Held, & Schrepp, 2008). Il comprend une évaluation de six dimensions (i.e., attractivité, efficacité, clarté, fiabilité, stimulation et nouveauté) à l'aide d'un ensemble de 26 questions. L'UEQ repose sur le modèle théorique d'Hassenzahl et emploie également un format de réponse sémantique différentielle (e.g., la paire de mots « lent – rapide »

est proposée pour évaluer la dimension nouveauté). Voss, Spangenberg et Grohmann (2003) ont également proposé l'échelle HED/UT permettant d'évaluer les qualités utilitaires et hédoniques (e.g., « fonctionnel / non fonctionnel » pour évaluer les qualités utilitaires).

Plus globalement, la majorité des études (e.g., Hassenzahl, 2004 ; Lallemand, Koenig, et al., 2015 ; Minge et al., 2016) s'intéressant à la structure factorielle des échelles UX emploie des Analyses en Composantes Principales (ACP – *Principal Component Analysis*). Cette méthode vise en premier lieu à transformer un ensemble de variables corrélées en un nouvel ensemble réduit de variables décorrélées (Abdi & Williams, 2010). Cette approche se distingue de l'Analyse Factorielle Exploratoire (*Exploratory Factor Analysis – EFA)* qui vise à identifier la structure latente des corrélations observées entre les variables mesurées. Ainsi, les ACPs visent à réduire le nombre de variables en expliquant la même quantité de variance alors que les EFAs visent à extraire les facteurs sous-jacents non mesurables directement (Suhr, 2005). Floyd et Widaman (1995) ont montré sur la base des simulations Monte Carlo de Hu, Bentler et Kano (1992) que les ACPs et les EFAs peuvent fournir des résultats très similaires mais uniquement dans des conditions d'utilisations optimales (e.g., échantillon de plus de 300 participants, au moins 10 items par facteur). Dans le cas inverse, les EFAs donneraient de meilleurs résultats. Ainsi, il serait nécessaire d'utiliser des ACPs uniquement lorsqu'on souhaite réduire le nombre de variables et des EFAs lorsqu'on fait l'hypothèse d'une structure factorielle sous-jacente comme dans le processus de création d'une échelle de mesure.

D'autres méthodes ont été proposées pour évaluer des composantes de l'UX comme par exemple la complétion de phrases (Kujala & Nurkka, 2012), qui consiste à proposer des phrases que le participant doit compléter par rapport à son jugement du système qu'il évalue, ou le test des 5 secondes (Gronier, 2016), qui revient à présenter pendant 5 secondes une interface puis à recueillir l'avis des participants sur celle-ci. Ces méthodes peuvent être des outils utiles pour les praticiens souhaitant pouvoir émettre des recommandations rapides sur les systèmes. Cependant, ces outils non standardisés ne permettent pas d'évaluer de manière systématique (notamment du fait que ce sont des méthodes qualitatives) les différentes dimensions semblant être incluses dans le jugement utilisateur.

5.5. Synthèse des modèles de l'UX

Comme il a été présenté précédemment, de nombreux modèles ont été proposés dans la littérature. Il existe plusieurs points communs et différences entre les divers modèles théoriques.

Premièrement, la majorité des modèles reconnaissent le rôle de trois facteurs sur l'UX : l'utilisateur, le contexte et le système. Dans le cadre du projet UXPEE dans lequel s'intègre cette thèse, il était nécessaire de répondre à la contrainte d'inclure uniquement des variables sur lequel le système évalué pouvait avoir une influence (e.g., facilité d'usage perçue, image de soi renvoyée par le produit). Ainsi, certaines variables (e.g., anxiété informatique ; Maricutoiu, 2014) ont été exclues des études bien qu'elles puissent avoir un effet sur le jugement envers un système.

Deuxièmement, la plupart des théories de l'UX s'accordent sur un découpage selon deux catégories de qualité d'un produit. La première catégorie correspondrait aux qualités instrumentales, pragmatiques du produit. C'est-à-dire à tous les éléments liés aux qualités fonctionnelles du produit. L'autre catégorie serait, quant à elle, composée de toutes les caractéristiques dites non fonctionnelles d'un système, aussi appelées qualités hédoniques ou non instrumentales. Ces éléments, nouvellement intégrés ou évoqués au sein des systèmes, ont pour rôle de dépasser les attentes fonctionnelles que les utilisateurs ont de leurs produits. Ces deux éléments seront pris en compte dans la suite des travaux de cette thèse.

Troisièmement, la sphère émotionnelle impliquée au cours de l'usage semble jouer un rôle important. Cette composante des modèles de l'UX est aujourd'hui la moins connue. Cependant, son rôle important sur la prise de décision et les comportements (e.g., Lerner et al., 2015) est largement reconnu. Ainsi, l'influence des émotions sur l'UX pourrait être prise en compte mais aussi, plus largement, les effets de plusieurs variables qui sont rattachées à la sphère affectivo-motivationnelle et qui ont été présentées dans le chapitre précédent.

Dernièrement, la norme ISO 9241-210 (2010) proposant une définition de l'UX met en avant la notion d'UX anticipée, c'est-à-dire d'un jugement envers un système au préalable de l'usage. Cependant, la plupart des précédentes études sont orientées sur le jugement après usage. Ainsi, les éventuelles différences entre le jugement avant

et après usage ne sont pas incluses dans les modèles théoriques et semblent un élément intéressant à étudier dans de futurs travaux.

En Résumé

Pour aller au-delà des simples qualités fonctionnelles des produits, la notion d'UX a été proposée. Ce concept vise à englober les qualités fonctionnelles et non fonctionnelles d'un produit ainsi que les réactions affectives qu'il provoque chez l'utilisateur. Néanmoins il n'existe pas aujourd'hui, à notre connaissance, d'outil permettant d'évaluer toutes ces notions simultanément ni d'analyse sur les structures sous-jacentes entre ces variables.

6. Modèles intégrateurs des notions UX et d'acceptabilité/acceptation

À notre connaissance, les modèles de l'UX et d'acceptabilité/acceptation n'ont été mis en lien qu'au sein d'un nombre réduit de recherches. En effet, ces modèles peuvent avoir des objectifs et objets d'études différents. De fait, les modèles de l'UX sont majoritairement centrés sur la perception des qualités fonctionnelles et non fonctionnelles d'un produit alors que les modèles de l'acceptabilité s'attachent à l'étude des convictions des individus (e.g., facilité d'usage perçue, utilité perçue) sur l'intention d'usage d'un produit. Or, il semble essentiel de mixer ces approches afin d'obtenir une représentation plus complète du jugement des utilisateurs. Ainsi, les travaux précédents sur les liens entre UX et acceptabilité/acceptation seront présentés puis un nouveau modèle intégrateur sera proposé. Ce dernier servira de base pour le développement de notre outil de mesure et sera évalué au sein de plusieurs études.

6.1. Modèle de van Schaik et collaborateurs

Van Schaik et Ling (2011) se sont intéressés à la mise en lien d'un modèle de l'UX (modèle Hassenzahl, 2005a) et d'un modèle de l'acceptation (TAM ; F. D. Davis, 1989). La mise en lien de ces deux modèles repose sur des travaux témoignant d'une relation entre esthétisme et intention d'usage (Cyr et al., 2006 ; van der Heijden, 2003), un effet positif du plaisir perçu (mesuré comme de la motivation intrinsèque) et de la facilité d'usage perçue (mesurée comme de la motivation extrinsèque) sur l'intention d'usage (F. D. Davis et al., 1992) ainsi qu'un lien entre les qualités hédoniques et à la fois le plaisir perçu et l'utilité perçue (e.g., Cyr et al., 2006 ; Overby & Lee, 2006).

Afin de tester leurs hypothèses, van Schaik et Ling ont présenté aux participants deux versions d'un site web (i.e., Wikipedia) où ils devaient soit chercher une information, soit simplement naviguer sur le site web. Sur la base de données collectées, les auteurs ont testé un modèle où les variables de l'UX sont les prédicteurs des variables du TAM. Ils ont obtenu les résultats présentés en figure 15.

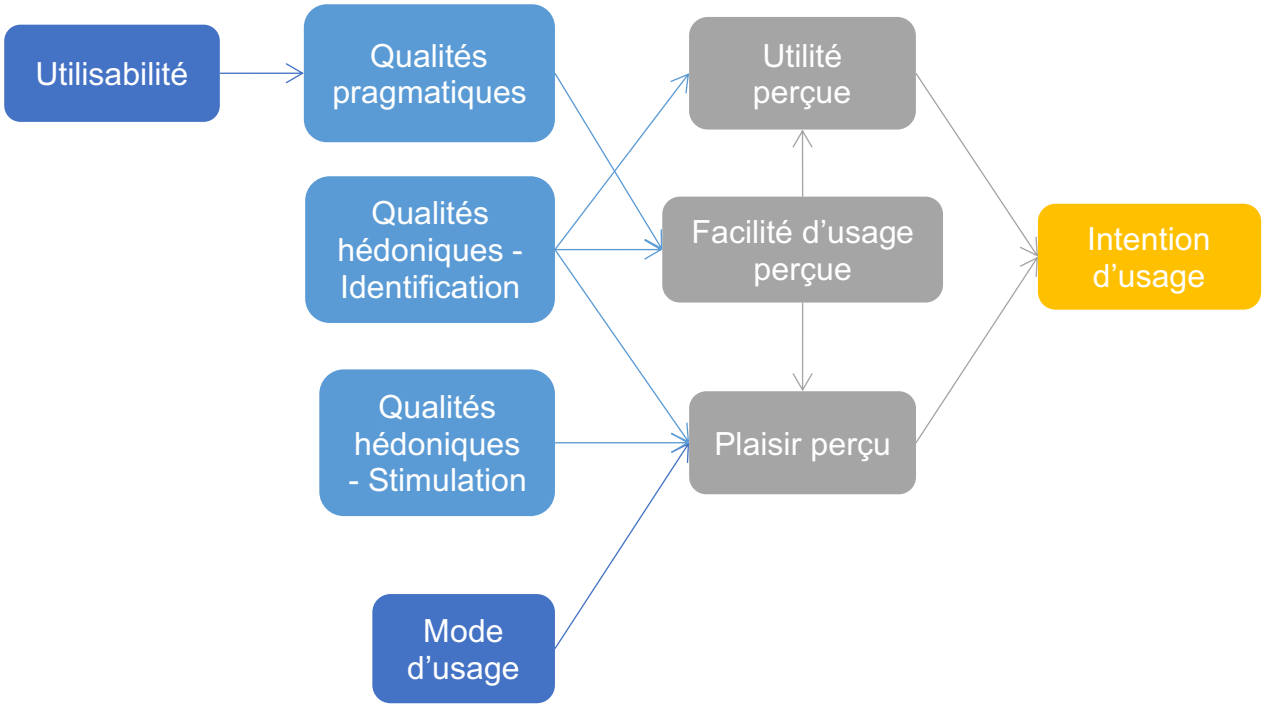


Figure 15. Modèle intégré de l'UX et de l'acceptabilité (van Schaik & Ling, 2011)[9, 10]

En conclusion, ce modèle semble un support intéressant pour de nouvelles recherches. Cependant, les émotions ne sont pas prises en compte. Il semble également pertinent d'évaluer ce modèle dans le cadre de l'utilisation d'un système davantage susceptible d'influencer des qualités non fonctionnelles (e.g., stimulation) qu'une page Wikipédia.

[9] Seuls les liens significatifs sont présentés.

[10] Régression des moindres carrés partiels (*Partial Least Squares*).

Plus récemment, ce modèle, auquel une variable émotion a été ajoutée, a servi de base à l'étude de sites d'informations (Aranyi & van Schaik, 2015a, 2015b) et le modèle suivant (voir figure 16) a été retenu à partir de leurs données.

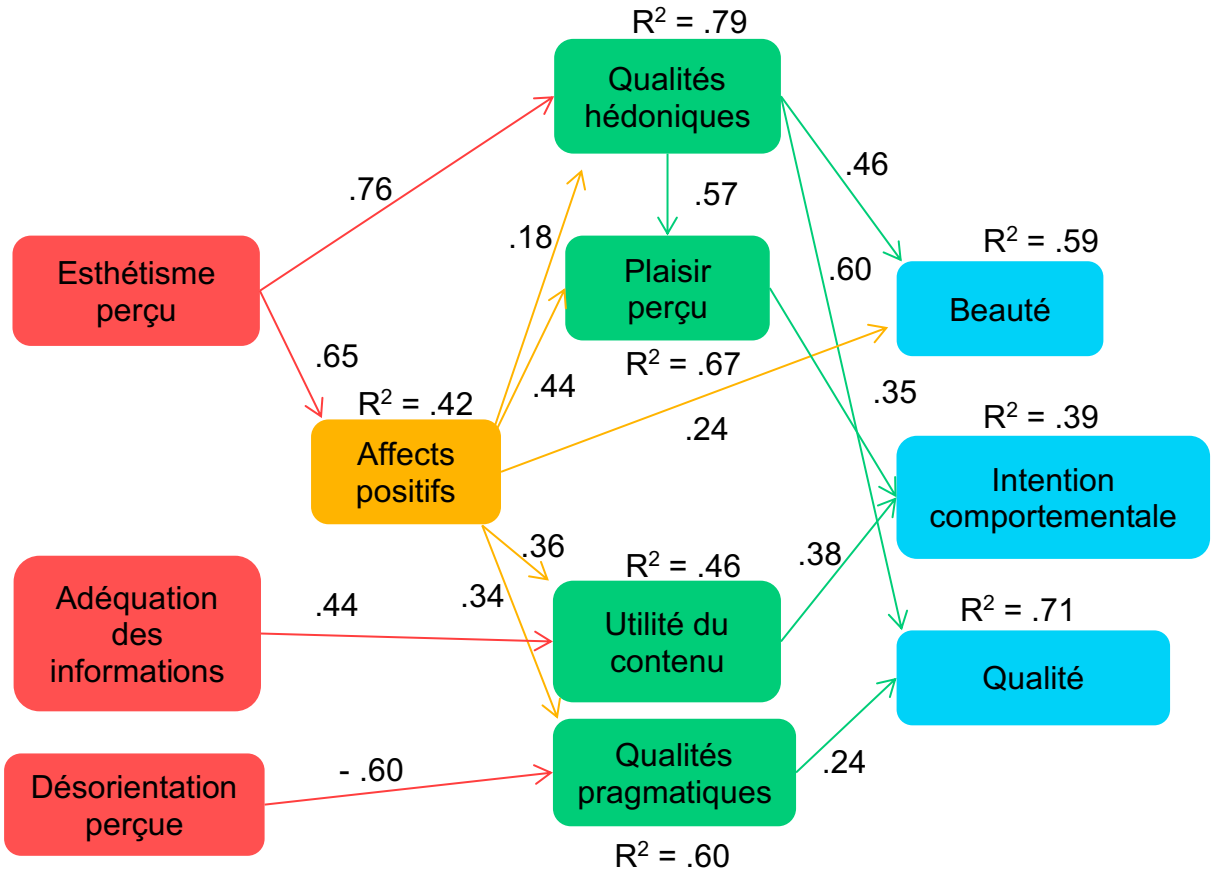


Figure 16. Modèle utilisé par Aranyi & van Schaik (2015a) pour l'étude de sites d'informations

Selon ce modèle, il y aurait une influence des affects positifs sur l'UX (i.e., qualités hédoniques et pragmatiques). Ces éléments influenceraient les évaluations de qualité et de beauté ainsi que l'intention d'usage. Ces résultats mériteraient d'être adaptés et testés sur d'autres systèmes que des sites d'informations. De plus, ce modèle n'intègre pas certaines variables pouvant avoir une influence sur le jugement (e.g., influence sociale, image de soi ; Venkatesh et al., 2012a).

6.2. Modèle de Février

Au cours de sa thèse, Février (2011) s'est notamment intéressée au rôle des affects dans la construction de l'intention d'usage envers un environnement numérique

de travail (ENT). Selon ces travaux, la difficulté (réussite/échec) aurait un effet sur les affects et les affects auraient une influence sur l'intention d'usage.

Deuxièmement, les résultats ont mis en avant une dynamique dans la perception et le jugement de l'interface au cours du temps. En effet, il a été montré que l'état affectif, l'utilisabilité perçue, les qualités affectives perçues, l'utilité perçue et l'effort mental perçu évoluaient au fur et à mesure de la réalisation des tâches et de la difficulté de ces dernières au sein d'une session d'expérience en laboratoire d'environ 1h.

Enfin, deux études ont été menées afin de mettre en lien des variables dites antécédentes (compétences en informatiques, anxiété informatique, etc.), des variables de perception (utilisabilité, qualités hédoniques, affects, etc.), des variables de jugements (utilité et attrait) et l'intention d'usage (Février, Rouxel, & Jamet, 2012). Les modèles suivants ont été mis en avant : voir figure 17 et figure 18.

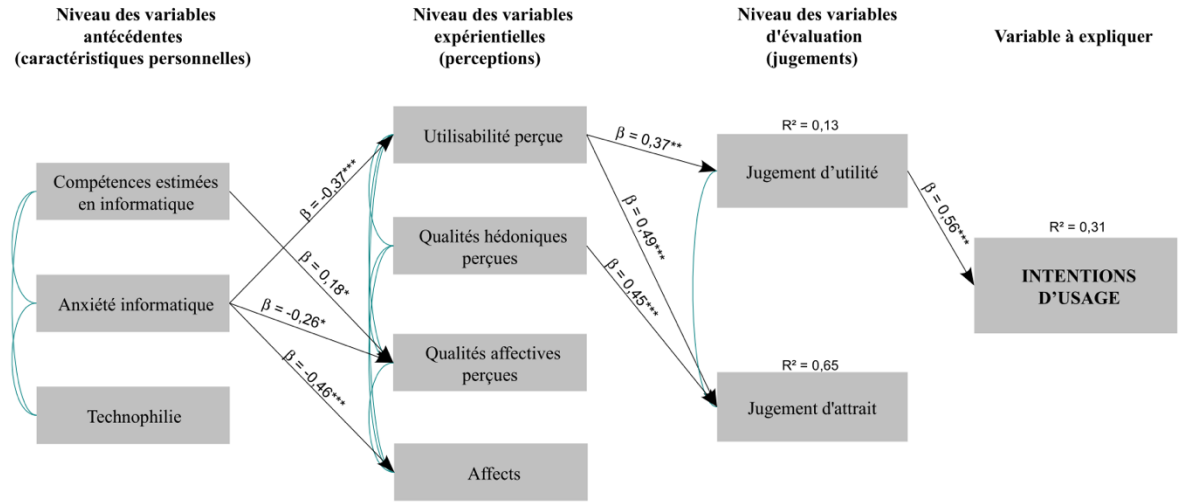


Figure 17. Modèle 1 testé par Février (2011)

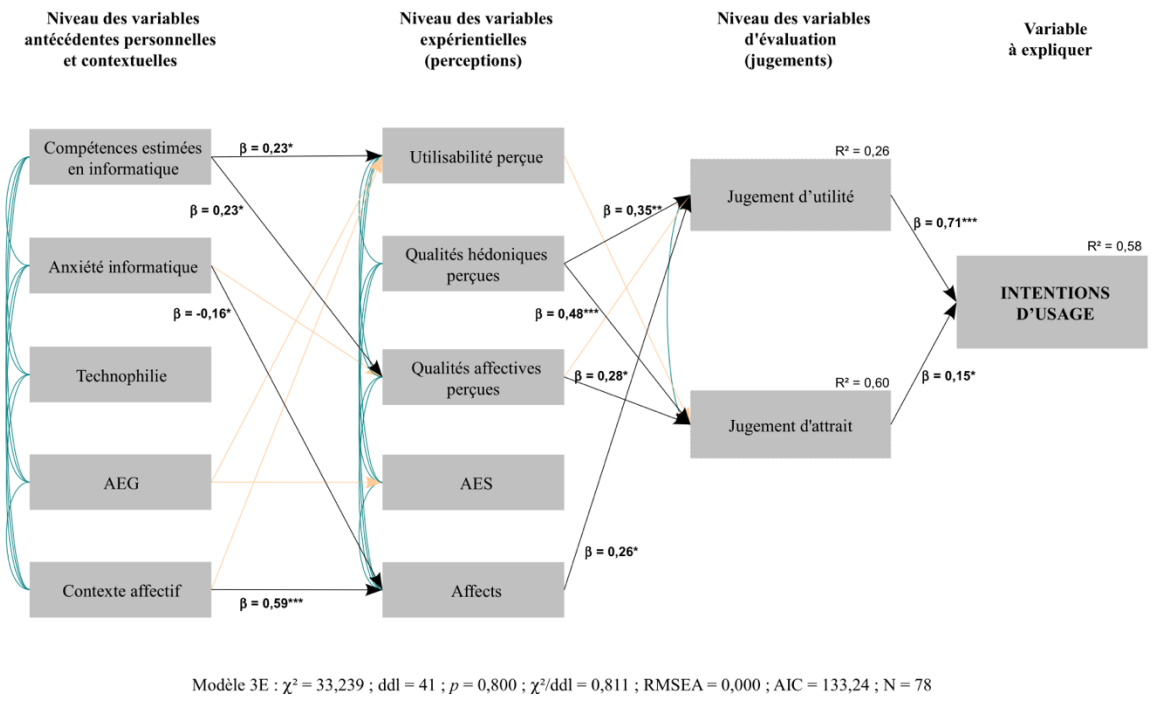


Figure 18. Modèle 2 testé par Février (2011)

Ces résultats montrent une influence des variables antécédentes sur la perception des qualités du système. Cette perception influencerait à son tour les jugements des utilisateurs. Finalement, ces jugements influencent l'intention d'usage. On note tout de même, comme pour les études sur l'acceptabilité, que l'utilité perçue joue un rôle prépondérant dans l'intention d'usage. Ces travaux offrent un modèle intégrant à la fois des aspects liés à la personnalité et des éléments de l'acceptabilité, en passant par des variables liées à l'expérience utilisateur.

En Résumé

Dans la littérature, plusieurs recherches ont été conduites pour intégrer les différentes approches théoriques (e.g., acceptabilité et UX). Néanmoins, il n'existe pas, à notre connaissance, de modèle intégrant à la fois les théories de l'acceptabilité, de l'UX et les variables affectivo-motivationnelles qu'il semble nécessaire de combiner pour obtenir une vue plus complète du jugement utilisateur.

7. Évolution du jugement

7.1. Évolution au cours du temps

Au-delà de la mise en lien entre l'UX et l'acceptabilité, il se développe depuis quelques années un intérêt pour les recherches longitudinales en IHM (Harbich & Hassenzahl, 2016 ; Karapanos, Jain, & Hassenzahl, 2012). Cependant, le nombre d'études reste faible (von Wilamowitz-Moellendorff, Hassenzahl, & Platz, 2006). Ainsi, encore aujourd'hui, l'évolution des variables liées à l'UX mais aussi leurs rôles au cours de l'usage sont encore mal connus. Parmi les recherches déjà effectuées, von Wilamowitz-Moellendorff et ses collaborateurs (2006) ont demandé à des utilisateurs d'évaluer rétrospectivement leur téléphone mobile au cours d'un entretien. En ce qui concerne l'évaluation des qualités instrumentales des téléphones, l'utilité perçue apparaît stable et l'utilisabilité perçue s'améliorerait au cours du temps. Pour les qualités hédoniques, ces propriétés seraient jugées plus négativement avec le temps. Les motifs de ces évolutions ont été répartis selon deux catégories : la familiarité avec le produit et la comparaison avec d'autres produits qui peuvent être, par exemple, plus récents et donc potentiellement avoir de meilleures qualités. D'autres travaux ont été réalisés dans la même approche ethnographique que celle de Karapanos (Karapanos, Martens, & Hassenzahl, 2012 ; Karapanos, Zimmerman, Forlizzi, & Martens, 2009, 2010 ; Kujala, Roto, Väänänen-Vainio-Mattila, Karapanos, & Sinnelä, 2011). Dans l'ensemble de ces recherches, la méthode consiste à interroger à plusieurs reprises des utilisateurs sur leur usage d'un produit. Pour cela, les auteurs se basent généralement sur une approche holistique de l'UX. Afin de justifier ce positionnement, plusieurs critiques sont rapportées à l'égard de l'approche multi-composantes (Karapanos, Zimmerman, et al., 2010). Premièrement, Karapanos et ses collègues considèrent que les modèles structuraux (utilisés pour évaluer les hypothèses théoriques) contraignent à considérer les construits latents et les liens entre les variables comme constants au cours du temps. Or, à juste titre, l'influence de certaines variables et leurs valeurs sont susceptibles de changer au fur et à mesure de l'usage. Cependant, il existe des modèles permettant de tenir compte de cette variabilité, comme les modèles de croissance latente (*Latent Growth Modeling*) (Hox & Stoel, 2005). Deuxièmement, ils reprochent à cette approche l'utilisation de questionnaires où les individus auraient des difficultés à se positionner. Pour cela ils se basent sur les

résultats de Schwarz (2007), ainsi que ceux de Larsen et ses collaborateurs (2008), montrant que les individus peuvent avoir des difficultés à répondre aux items. Cependant, les méthodes d'entretiens comportent également des défauts : il est communément considéré que les résultats obtenus, par la nature même des techniques d'entretiens, sont difficilement généralisables (Harris & Brown, 2010).

Cependant, leurs travaux offrent des résultats intéressants. Ils ont ainsi constaté trois phases dans l'adoption d'un produit : l'orientation, l'incorporation et l'identification (voir figure 19).

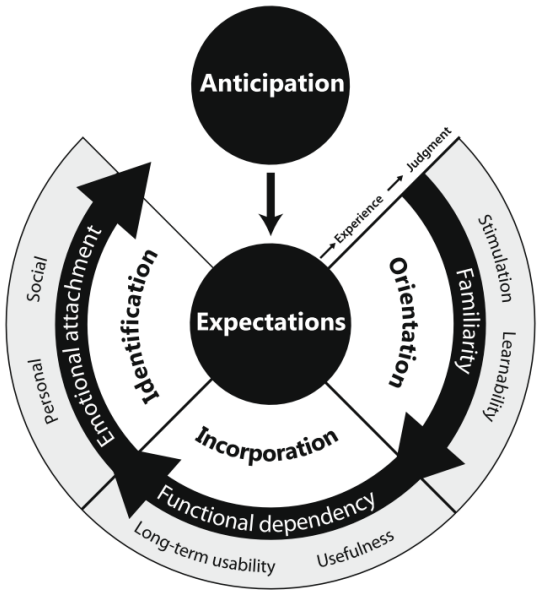


Figure 19. Modèle proposé par Karapanos & al. (2009)

À chacune de ces phases, des qualités particulières du produit sont associées. Ainsi, la phase d'orientation serait rattachée à la stimulation et à l'apprenabilité du produit. La deuxième (l'incorporation) est consistuée de l'utilité et l'utilisabilité perçue du système. Enfin, la troisième et dernière phase (l'identification) est liée à la personnalisation et à l'aspect social du produit. Ces résultats apportent des premiers éléments sur l'évolution de l'UX au cours de l'usage et méritent d'être étudiés au sein d'une approche multi-composante de l'UX.

De plus, l'évolution de l'UX au cours d'une session d'usage a également fait l'objet de diverses études. Ainsi, le jugement de l'UX serait hautement dynamique et pourrait évoluer au cours d'une seule et même interaction (Wimmer, Wöckl, Leitner, & Tscheligi, 2010). Plus récemment, Harbich & Hassenzahl (2016) se sont intéressés à

l'évolution de l'UX au cours du temps dans un contexte professionnel en présentant notamment de nouvelles méthodes d'analyses pour traiter ce type de données. Enfin, van Schaik, Zeng, & Spears (2014) ont proposé une nouvelle approche dans l'étude de l'évolution au cours du temps de l'UX (et des variables qui y sont rattachées) reposant sur les réseaux bayésiens dynamiques. Cette technique vise à découvrir les paramètres sous-jacents régissant les liens entre les variables au cours du temps (Mitra, Zaslavsky, & Åhlund, 2011). Ainsi, il serait théoriquement possible d'estimer à l'avance l'évolution des paramètres sur les différentes variables à partir de données récoltées au début de l'usage.

En conclusion, un nombre limité de publications a été dédié à l'étude de l'évolution au cours du temps de l'UX, notamment lié aux difficultés de conduire de telles études (Harbich & Hassenzahl, 2016). On peut également noter qu'à l'exception des travaux récents de Harbich & Hassenzahl (2016) dans un contexte professionnel, ces recherches reposent sur une approche qualitative.

7.2. Outils de mesure de l'évolution au cours du temps

Pour mesurer l'évolution de l'expérience utilisateur au cours du temps, plusieurs approches sont proposées. Karapanos, Martens et Hassenzahl (2010) relèvent quatre approches : l'évaluation transversale (i.e., évaluation avec des échantillons différents utilisant le système depuis des durées différentes), l'évaluation par mesures répétées (i.e., les mêmes participants évaluent plusieurs fois le même système à des temps d'usage différents), l'évaluation rétrospective (i.e., les participants doivent évaluer rétrospectivement le système) et l'évaluation longitudinale (i.e., évaluation sur une longue période par les mêmes participants du même système). Chacune de ces approches peut reposer sur des outils différents. On retrouve par exemple, la méthode des questionnaires (e.g., Harbich & Hassenzahl, 2016), les courbes d'évaluation UX (Karapanos, Martens, et al., 2012 ; Kujala et al., 2011) ou encore les journaux de bord (e.g., Day Reconstruction Method ; Karapanos et al., 2009).

En Résumé

La dimension temporelle dans le jugement utilisateur a connu un engouement ces dernières années. Cependant, la majorité des approches de la littérature est basée sur des résultats qualitatifs.

8. Approche proposée

8.1. Les éléments essentiels

Au regard de l'ensemble des travaux présentés précédemment, la compréhension du jugement utilisateur requière la mise en lien de nombreuses théories entre elles et par conséquent la prise en compte de nombreuses variables. Ainsi, trois éléments différents (voir figure 20) semblent jouer un rôle : les antécédents, les perceptions du système et les conséquences.

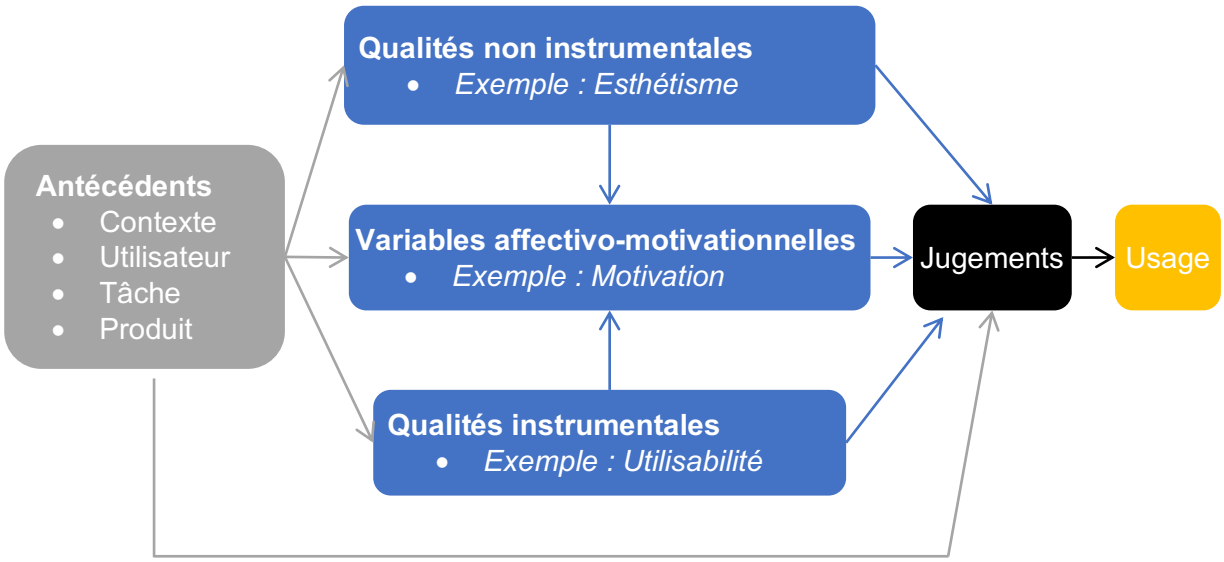


Figure 20. Les composantes du jugement utilisateur envers un produit technologique : Antécédents, Perceptions et Conséquences

Les antécédents à l'usage correspondent aux variables qui peuvent influencer l'ensemble des liens et des autres éléments du modèle. On y retrouve les caractéristiques de l'individu (e.g., traits de personnalité, expérience, etc.), le contexte (notamment professionnel et loisirs), la tâche et le produit. Ces variables ne seront pas étudiées dans cette thèse. En effet, bien que les effets de ces variables soient reconnus au sein de nombreuses études (e.g., Djamasbi, Strong, & Dishaw, 2010 ; Venkatesh et al., 2003), la modification du produit n'aura pas d'influence sur ces variables et par conséquent cela n'est pas considéré comme approprié au regard des contraintes imposées par le projet UXPPE.

Le second niveau correspond aux perceptions des utilisateurs. Il est composé de la perception des qualités fonctionnelles et non fonctionnelles du produit mais également de l'influence des processus liés à la sphère affectivo-motivationnelle.

Enfin, le dernier niveau correspond aux conséquences des étapes présentées précédemment. Cela correspond à la formation d'un jugement (e.g., évaluation globale, intention d'usage) à partir des variables présentées ci-dessous qui influenceraient l'usage du produit.

8.2. Modèle théorique intégratif

Au sein de la littérature, de nombreuses théories et modèles ont été développés pour comprendre le jugement utilisateur. Il semble pertinent de tenter d'incorporer ces différents éléments théoriques au sein d'un modèle intégratif afin de saisir au mieux le jugement utilisateur envers un produit technologique. Ainsi, le modèle proposé vise tout d'abord à prendre en compte les éléments liés au jugement avant usage (voir Figure 21). En effet, le jugement avant usage et ses composantes devraient influencer le passage au premier usage et donc l'adoption du produit à plus long terme. Cette influence devrait diminuer au fur à mesure des usages. On peut noter que la majorité des précédents travaux a ignoré le jugement avant usage (Šumak & Šorgo, 2016) ce qui conforte l'intérêt pour cet élément. De plus, au regard de la littérature, des évolutions devraient apparaitre sur les variables liées à l'UX au fur et à mesure de l'usage (voir Figure 21).

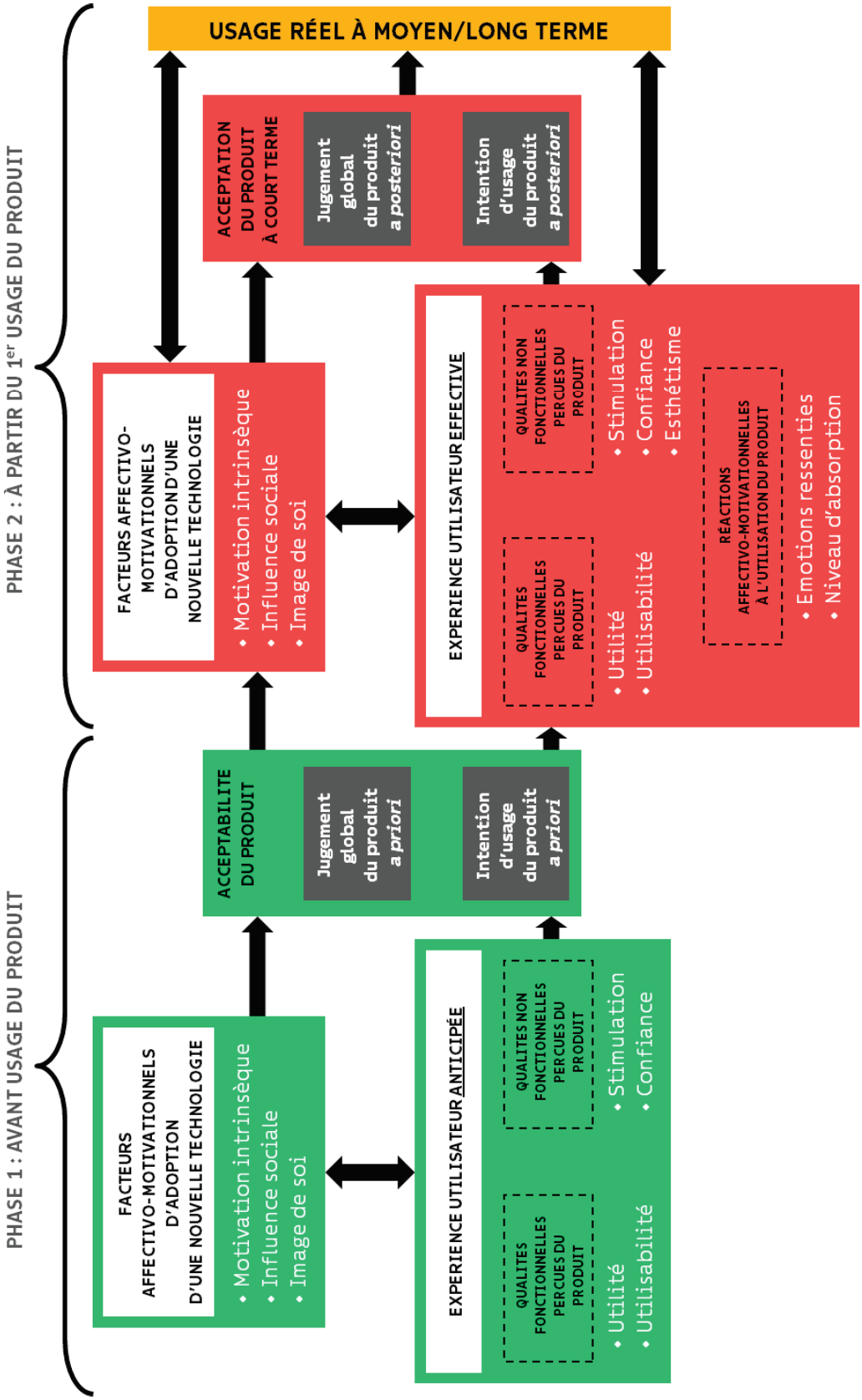


Figure 21. Modèle intégratif et dynamique de l'acceptabilité/acceptation d'une nouvelle technologie

9. Conclusion

9.1. Une variété de modèles existants

Au sein de cet état de l'art, de nombreux modèles et études sur l'évaluation des produits technologiques par les utilisateurs ou potentiels utilisateurs ont été présentés. Ces travaux sont principalement rattachés aux champs théoriques de l'acceptabilité/acceptation et de l'UX.

En résumé, le champ théorique de l'acceptabilité et de l'acceptation est orienté sur l'identification des facteurs à l'origine de l'adoption ou du rejet d'une technologie (e.g., Venkatesh et al., 2012a). Cette approche, la plus ancienne, a d'abord mis en avant l'influence de variables liées aux facteurs fonctionnels des produits (e.g., utilité perçue ; F. D. Davis, 1989) puis des facteurs moins techniques (e.g., confiance perçue ; Flavián, Guinalíu, & Gurrea, 2006) au fur et à mesure des avancées technologiques et des travaux de recherche. En ce qui concerne le champ théorique de l'UX, il est centré sur la mise en lumière des facteurs influençant le jugement global envers un système (e.g., Hassenzahl, 2004). Ces travaux sont nés de la mise en lumière de la perception de qualités non fonctionnelles (e.g., stimulation ; Hassenzahl & Monk, 2010) davantage en lien avec l'humain qu'avec la technologie.

Les différents travaux, notamment ceux réalisés dans le champ de l'acceptabilité et de l'UX, ont conduit à une multiplication de modèles et la découverte de nombreuses variables. Cependant, cette multiplication de publications ne signifie pas forcément une meilleure compréhension du ressenti des utilisateurs lors de l'usage d'un produit. En effet, il est apparu qu'un certain nombre de ces variables aux appellations diverses se recouvrent sur le plan théorique au point, parfois, d'être mesurées de manière identique (e.g., plaisir perçu et motivation intrinsèque). De plus, on peut noter que les champs théoriques de l'acceptabilité et de l'UX ont très rarement été mis en lien (van Schaik & Ling, 2011). Or, ils semblent complémentaires. Intégrer ces deux approches au sein d'un modèle unifié pourrait permettre d'obtenir une compréhension plus profonde de l'adoption des produits technologiques par les utilisateurs. Selon Obrist et collaborateurs (2012), de nouveaux travaux doivent être conduits autour de la théorie et la méthodologie dans l'UX. Law et van Schaik (2010)

proposent notamment d'utiliser de nouvelles méthodes statistiques pour tester ces modèles.

9.2. Problématique de la thèse

Comme présenté en préambule, cette thèse s'inscrit au sein du projet UXPPE visant notamment à développer de nouveaux travaux et outils pour l'évaluation des produits technologiques. Ainsi, la problématique générale de cette thèse est la suivante : quelles variables influencent le jugement utilisateur avant et après usage ? Cette problématique englobe trois questions sous-jacentes. Tout d'abord, quels sont, dans la littérature, les modèles les plus pertinents et leurs variables rattachées pour comprendre le jugement utilisateur ? Cette première partie vise à la conception d'un modèle unifiant les variables et théories présentes dans la littérature. Deuxièmement, comment évaluer ces variables considérées comme les plus appropriées ? Enfin, comment ces variables évoluent au cours du temps ?

Au sein de ce document, les études sur le jugement utilisateur seront présentées séquentiellement, avec, tout d'abord, les travaux sur le jugement avant usage, puis sur le jugement après usage, puis l'évolution entre le jugement avant et après usage, et enfin des travaux sur la mise en place d'outils pour les praticiens.

Chapitre 1 : Modélisation de l'adoption avant usage[11]

L'adoption (i.e., utilisation ou rejet du système par les individus) d'un nouveau produit est un long processus. Généralement, les individus commencent par se construire des représentations et des attentes (i.e., « croyances […] sur un produit constituées avant son utilisation réelle » [notre traduction] ; Raita & Oulasvirta, 2011) à propos de différents systèmes offrant des services similaires en se basant sur plusieurs critères (Hsu & Lin, 2015). Ensuite, à partir de ces représentations, les individus peuvent comparer ces différents produits, potentiellement en sélectionner un et commencer à l'utiliser.

Communément, les personnes construisent ces représentations et réalisent leur choix sans avoir utilisé le produit. Par exemple, lors du choix d'une application mobile au sein d'un magasin d'applications (e.g., Google Play), l'utilisateur ne peut tester au préalable toutes les applications avant de les télécharger (d'autant plus si parmi les choix possibles il y a des applications payantes). Il doit donc réaliser son choix, en fonction de certains critères. Ainsi, comprendre les facteurs qui conduisent à l'adoption par les utilisateurs d'un système apparaît comme un champ de recherche essentiel afin de construire des outils pertinents pour l'évaluation des systèmes. A ce titre, de multiples travaux ont été conduits depuis de nombreuses années dans le champ des Interactions Homme-Machine. Néanmoins, les précédentes recherches ont été majoritairement orientées sur le jugement après usage (e.g., Venkatesh et al., 2003). Les représentations avant usage, notamment les facteurs qui les composent, ont rarement été étudiées (Karahanna et al., 1999 ; Šumak & Šorgo, 2016 ; S. ang et al., 2012). Or, les représentations avant usage constituent l'un des éléments essentiels pour appréhender au mieux les processus de décision chez les utilisateurs.

[11] Une partie des résultats concernant cette étude a été publiée dans la conférence Interaction-Homme Machine (IHM'15) à Toulouse (Martin, Erhel, Jamet, & Rouxel, 2015). Un article journal (*Which factors influence behavioral intention before use? Studies with two mobile applications using structural equation modeling*) a également été soumis.

Au sein de ce chapitre, deux études ont été conduites pour comprendre avec davantage de précision les facteurs influençant le jugement utilisateur envers un système avant usage en se basant sur les champs théoriques de l'acceptation et de l'expérience utilisateur (UX). Dans ce cadre, il a été choisi d'étudier exclusivement le jugement des utilisateurs envers des applications mobiles, et ce principalement pour des raisons pratiques. En effet, la sélection des applications par les individus représente un terrain d'expérimentation relativement écologique pour étudier le jugement avant usage. Dans un contexte d'usage classique, les utilisateurs peuvent être amenés à choisir des applications mobiles sans avoir la possibilité de les utiliser préalablement. De plus, il est relativement aisé de trouver un large échantillon de participants possédant un smartphone en raison de leur large déploiement.

1. Acceptabilité / Acceptation et expérience utilisateur : jugement avant et après usage

De nombreuses études ont été réalisées pour mieux appréhender l'adoption des produits. Elles sont globalement reliées à deux champs théoriques. Le premier est composé des travaux sur l'acceptabilité et l'acceptation. Il vise à identifier les facteurs facilitateurs ou inhibiteurs de l'adoption de produit(s) technologique(s) (e.g., Dillon & Morris, 1996). Même si l'acceptabilité et l'acceptation sont couramment utilisées comme synonymes, l'acceptabilité se définit comme le jugement prospectif envers une technologie avant usage (e.g., Schade & Schlag, 2003). L'acceptation, quant à elle, peut être vue comme le jugement et les réactions comportementales envers un produit après usage (Schuitema et al., 2010). Le second champ de recherche, l'expérience utilisateur (UX), est centré sur les facteurs qui composent le jugement utilisateur envers un produit. Selon la norme ISO 9241-210 (2010), l'UX est définie comme représentant les « perceptions et réactions d'une personne qui résultent de l'utilisation effective et/ou anticipée d'un produit, système ou service ». Cependant, même si la distinction entre le jugement avant et après usage est mise en avant au sein de ces deux champs théoriques, les précédentes études sont largement centrées sur le jugement après usage (Karahanna et al., 1999 ; Šumak & Šorgo, 2016 ; S. ang et al., 2012). Ainsi, les facteurs composant le jugement avant usage ont rarement été explorés (ogasara, Popovic, Kraal, & Chamorro-Koc, 2011).

Au sein de ce chapitre, les facteurs du jugement après usage, inhérents aux champs théoriques de l'acceptation et de l'UX, ont été utilisés afin de modéliser le jugement avant usage. Pour évaluer l'influence de ces variables, plusieurs échelles ont, tout d'abord, été construites puis testées sur le plan psychométrique. Ensuite, les données collectées avec ces échelles ont été utilisées pour estimer l'effets des variables mesurées à l'aide de modèles à équations structurelles (*Structural Equation Modeling* – SEM). Pour modéliser le jugement utilisateur, il convient de préciser davantage le contexte théorique dans lequel s'inscrivent nos études.

2. Rappel du contexte théorique et hypothèses

À notre connaissance, la modélisation du jugement avant usage a été explorée au sein de peu d'études. Paradoxalement, la plupart des travaux sur l'avant usage ont collecté leurs données uniquement après l'utilisation du système (Karahanna et al., 1999 ; Šumak & Šorgo, 2016 ; S. ang et al., 2012). Au sein de deux de nos études, le jugement avant usage a été modélisé à partir de données collectées avant toute utilisation du produit. De cette façon, une meilleure compréhension de l'évolution du jugement au travers des usages devrait être possible.

2.1. Acceptabilité et Acceptation

L'adoption des produits a généralement été étudiée à travers le prisme du concept de l'acceptation. Le *Technology Acceptance Model* (TAM) (F. D. Davis, 1989) a largement été utilisé dans ce but (King & He, 2006). Plus récemment, un nouveau modèle incluant des variables de 8 modèles (dont le TAM) a été proposé, the *Unified Theory Of Acceptance And Use Of Technology* (UTAUT) (Venkatesh et al., 2003) ainsi que son extension, l'UTAUT2 (Venkatesh et al., 2012a). Cependant, même si l'UTAUT2 est centré sur le jugement après usage (i.e., acceptation), il semble représenter un cadre théorique intéressant pour étudier le jugement avant usage. En effet, certaines variables semblent être pertinentes pour évaluation avant et après usage.

L'UTAUT2 est composé des variables suivantes : Attentes de Performance, Attentes d'Effort, Influence Sociale, Motivation Hédonique, Conditions Facilitatrices, Prix et Habitudes[12].

Les Attentes de Performance (*Performance Expectancy* – PE) sont définies comme étant « le degré selon lequel un individu pense qu'un système pourrait l'aider à améliorer son travail » [notre traduction] (Venkatesh et al., 2003, p. 447). Également nommée Utilité Perçue (« la probabilité subjective et prospective qu'utiliser une application spécifique puisse améliorer son travail dans un contexte organisationnel » [notre traduction] ; F. D. Davis, 1989, p. 320), cette variable aurait un effet positif important sur l'acceptation selon plusieurs méta-analyses (Dwivedi et al., 2011 ; King & He, 2006 ; Ma & Liu, 2004 ; Schepers & Wetzels, 2007). Au sein des précédentes études, l'Utilité Perçue a seulement été évaluée après usage. Cependant, cette perception, reposant sur des descriptions ou expériences similaires, devrait également avoir une influence importante sur l'intention comportementale avant usage.

Hypothèse 1. L'Utilité Perçue devrait avoir une influence positive sur l'intention comportementale avant usage.

Parallèlement à l'Utilité Perçue, le concept d'Attentes d'Effort (*Effort Expectancy* - EE), aussi nommé Facilité d'Usage Perçue peut être défini comme « le degré de facilité associé à l'utilisation d'un système » [notre traduction] (Venkatesh et al., 2003, p. 450). Son influence sur l'intention d'usage a également été démontrée au sein de plusieurs méta-analyses (Dwivedi et al., 2011 ; Khechine et al., 2016 ; King & He, 2006 ; Ma & Liu, 2004 ; Schepers & Wetzels, 2007). Dans le contexte du jugement avant usage, l'effet de l'Utilité Perçue devrait exister mais avec probablement une influence plus faible que dans le contexte après usage. En effet, l'usage aurait un effet sur l'Utilisabilité Perçue (Venkatesh & Davis, 1996).

Hypothèse 2. L'Utilisabilité Perçue devrait avoir une influence positive sur l'intention comportementale avant usage.

[12] La version originale de l'UTAUT contient plusieurs variables modératrices (e.g., âge, sexe). Cependant très peu d'études reposant sur l'UTAUT ont inclus ces variables modératrices (Dwivedi, Rana, Chen, & Williams, 2011 ; Khechine, Lakhal, & Ndjambou, 2016). Par conséquent, elles ne seront pas prises en compte dans le présent chapitre.

Conjointement à ces qualités fonctionnelles, deux variables semblent pertinentes pour le jugement avant usage : l'Influence Sociale et la Motivation Hédonique.

L'Influence Sociale (*Social Influence* – SI) est définie comme étant « le degré selon lequel un individu perçoit que les personnes qui lui sont importantes croient qu'il ou elle devrait utiliser le nouveau système » [notre traduction] (Venkatesh et al., 2003, p. 451). Plusieurs méta-analyses sur de précédentes études ont montré une influence positive de cette variable sur l'intention comportementale (Dwivedi et al., 2011 ; Khechine et al., 2016 ; Schepers & Wetzels, 2007 ; Taiwo & Downe, 2013). Cette variable semble être uniquement influencée par l'entourage et non par l'usage. Dans les circonstances d'un jugement avant usage, l'Influence Sociale devrait avoir une influence sur l'intention d'usage.

Hypothèse 3. L'Influence Sociale devrait avoir une influence positive sur l'intention comportementale avant usage.

La Motivation Hédonique (*Hedonic Motivation* – HM), quant à elle, est définie comme étant « le fun ou le plaisir dû à l'utilisation d'une technologie » [notre traduction] (Brown & Venkatesh, 2005). Cette variable est semblable à la Motivation Intrinsèque (« le fait de réaliser une activité pour des raisons de satisfaction inhérentes à l'action plutôt que pour d'autres conséquences séparables » [notre traduction] ; R. M. Ryan & Deci, 2000) et son influence positive a été démontrée dans plusieurs études (e.g., Koufaris, 2002 ; Lee et al., 2005 ; Venkatesh et al., 2012a). Même si la Motivation Intrinsèque dépend des conséquences anticipées de la réalisation d'une activité, ces conséquences pourraient également être perçues par les individus sans aucune utilisation du produit.

Hypothèse 4. La Motivation Intrinsèque devrait avoir une influence positive sur l'intention comportementale avant usage.

Les Conditions Facilitatrices (*Facilitating Conditions* – FC) sont définies comme étant « le degré selon lequel un individu pense que son organisation et infrastructure technique peuvent l'aider dans l'usage du système » [notre traduction] (Venkatesh et al., 2003, p. 453). Au sein de nos études, ce prédicteur a été exclu pour deux raisons : 1) dans la version initiale de l'UTAUT, l'effet direct des FC sur l'intention d'usage n'a pas été confirmé ; 2) aucun support ne semble requis pour utiliser une application

mobile. Par conséquent, les FC ont été considérées comme non pertinentes et ne seront pas incluses dans les études suivantes.

La Valeur de Prix (*Price Value* – PV) est définie comme étant « le compromis cognitif des consommateurs entre les avantages perçus des applications et le coût monétaire pour les utiliser » [notre traduction] (Dodds et al., 1991). La majorité des applications mobiles sont aujourd'hui acquises gratuitement (Fu et al., 2013) et ainsi que celles sélectionnées dans les présentes études. Par conséquent la PV ne semble pas pertinente dans le présent contexte. Ainsi, la PV ne sera pas incluse dans les études suivantes.

Enfin, les Habitudes (*Habits*) sont définies comme « le construit perceptif qui reflète les résultats d'expériences antérieures » [notre traduction] (Venkatesh et al., 2012a). Les items utilisés pour évaluer cette dimension (e.g., l'utilisation de l'internet mobile est devenue une habitude pour moi ; Venkatesh et al., 2012a) sont directement influencés par l'usage même du produit. Par conséquent, évaluer cette variable avant usage ne semble pas pertinent et ne sera pas inclus dans les études à venir.

En résumé, quatre prédicteurs de l'intention d'usage provenant du champ de l'acceptation semblent pertinents dans le cadre de nos études : l'Utilité Perçue, l'Utilisabilité Perçue, l'Influence Sociale et la Motivation Intrinsèque.

2.2. Expérience utilisateur (UX)

Ces dernières années, le champ de l'expérience utilisateur (UX) s'est développé en réponse aux études montrant que les individus perçoivent les produits au-delà de leurs qualités fonctionnelles (Bargas-Avila & Hornbaek, 2011). Le modèle de l'UX d'Hassenzahl (Hassenzahl, 2004) apparait comme un des modèles les plus utilisés (Lallemand, Gronier, et al., 2015). Selon ce modèle, les individus percevraient les produits au travers de deux types de qualité : les qualités pragmatiques et les qualités hédoniques.

Les qualités pragmatiques sont définies comme les capacités perçues d'un produit pour atteindre un but (Hassenzahl et al., 2010). Ces qualités, correspondant à l'Utilité Perçue et l'Utilisabilité Perçue dans les modèles de l'acceptation, auraient un effet positif sur le jugement envers un produit après usage (e.g., Aranyi & van Schaik,

2015b ; Hassenzahl, 2004 ; Mahlke, 2007 ; Mahlke, Lemke, & Thüring, 2007 ; van Schaik & Ling, 2011). Ainsi, les qualités pragmatiques perçues d'un système sont évaluées à travers l'Utilité et l'Utilisabilité Perçues.

Les qualités hédoniques, quant à elles, sont reliées aux études mettant en avant le fait que les utilisateurs perçoivent les produits au-delà des qualités fonctionnelles (Bargas-Avila & Hornbaek, 2011). Ainsi, les qualités hédoniques seraient reliées à l'utilisateur lui-même (Hassenzahl, 2004). Selon le modèle d'Hassenzahl, la stimulation et l'identification[13] auraient un effet sur le jugement envers un produit.

La stimulation est reliée au besoin de challenge de nouveauté, prérequis pour le développement personnel (Hassenzahl, 2004, 2005a). Les études antérieures témoignent d'un effet positif sur le jugement (Aranyi & van Schaik, 2015a ; van Schaik & Ling, 2008). De plus, la satisfaction des besoins est un élément important dans la conduite des comportements (Hagger, 2006). La stimulation est rattachée aux fonctions fournies par le produit et peut être évaluée à partir des descriptions de l'application. Ainsi, ces attentes peuvent influencer l'intention comportementale avant usage.

Hypothèse 5. La Stimulation devrait avoir une influence positive sur l'intention comportementale avant usage.

La seconde variable proposée par Hassenzahl (2004), l'identification, est définie comme l'Image de Soi que le produit reflète de soi-même (Hassenzahl, 2004) et est reliée au besoin d'être socialement reconnu (Hassenzahl, 2005a). Selon les études précédentes sur le jugement après usage, l'identification aurait une influence positive sur l'intention d'usage (J. D. Jackson, i, & Park, 2013 ; Kwon & Chon, 2009). Les représentations à propos d'un produit peuvent être construites à partir des descriptions du produit et/ou les expériences préalables et devraient influencer l'intention comportementale avant usage.

[13] Dans Hassenzahl (2005a), une troisième dimension (évocation) a été inclus dans le modèle. Mais Hassenzahl (2004) indique que cette variable n'est pertinente que si le système « est en mesure de provoquer des souvenirs et agir comme un symbole du passé ». Ainsi, dans les présentes études, cette dimension ne semble pas pertinente et par conséquent ne sera pas incluse.

Hypothèse 6. L'Image de Soi devrait avoir une influence positive sur l'intention comportementale avant usage.

Parallèlement, la Confiance perçue semble également un facteur important dans l'adoption d'un produit technologique (Riegelsberger, Sasse, & McCarthy, 2005). Cette influence a largement été étudiée dans le contexte des sites web et un effet positif de cette variable sur l'intention d'usage a été démontré au sein de plusieurs études (Corbitt et al., 2003 ; Gefen et al., 2003 ; Jarvenpaa et al., 2000 ; Pavlou, 2003). Les représentations en termes de confiance peuvent exister avant d'avoir utilisé le produit. En effet, les expériences passées et/ou les descriptions du produit devraient conduire à une attente en termes de confiance et un effet de cette variable peut être attendu.

Hypothèse 7. La Confiance perçue devrait avoir une influence positive sur l'intention comportementale avant usage.

En conclusion, les choix effectués ici constituent une synthèse de l'UTAUT2 et du modèle d'Hassenzahl. L'UTAUT2 semble le modèle le plus courant pour décrire l'acceptation et semble ainsi pertinent (e.g., Šumak & Šorgo, 2016) pour modéliser l'acceptabilité (i.e., jugement avant usage). L'UX et notamment l'UX anticipée sont incluses au sein de la norme ISO 9241-210 (2010). Cependant, à notre connaissance, ce jugement avant usage n'a jamais été étudié au sein d'une approche quantitative. Par conséquent, la combinaison de l'UTAUT2 et des variables d'Hassenzahl semble représenter une approche intéressante pour modéliser le jugement avant usage. Afin de prédire le jugement avant usage avec ces variables, deux études ont été conduites avec deux échantillons indépendants de volontaires. La conduite de deux études indépendantes doit permettre de s'assurer de la réplicabilité des résultats. Des échelles de mesure ont été spécifiquement construites et testés au sein de ces études. L'influence des variables évaluées a été estimée avec des modèles à équations structurelles (SEM), comme recommandé par Law et van Schaik (2010). En effet, les SEMs permettent de tester toutes les hypothèses au sein de la même modélisation en estimant simultanément tous les paramètres (Lei & Wu, 2007) et, par conséquent, l'effet des variables.

3. Étude 1 : Application Sleep Better

Une première étude a été conduite pour tester l'effet des variables reliées aux champs théoriques de l'acceptation et de l'UX sur l'intention comportementale avant usage. Ainsi, les participants ont évalué une application mobile qu'ils n'avaient jamais utilisée, au cours d'une étude en ligne.

3.1. Méthode

3.1.1. Matériel

Les participants ont évalué l'application Sleep Better, une application de suivi du sommeil. Cette application a été sélectionnée parce qu'elle est simple, gratuite, faiblement distribuée[14] et potentiellement ludique. Avant de compléter les échelles, une courte présentation de l'application était proposée aux participants (puisqu'ils n'ont jamais utilisé l'application). Cette présentation correspond typiquement à la description proposée par les fabricants sur les magasins d'applications (i.e., AppStore et Google Play). Elle est composée d'un paragraphe textuel décrivant les fonctions de l'application ainsi que plusieurs captures d'écran de l'application : voir ci-dessous (présentation complète en Annexe 7).



Figure 22. Captures d'écran de l'application Sleep Better

[14] Pour évaluer les représentations et jugements avant usage, il est nécessaire que les participants n'aient jamais utilisé le système. Ainsi, il est essentiel de sélectionner un produit potentiellement peu utilisé pour faciliter le recrutement.

« Suivez votre sommeil, analysez vos rêves, améliorez vos heures de coucher et réveillez-vous mieux que jamais avec Sleep Better de Runtastic ! Vous avez des difficultés à vous endormir ? Vous voulez découvrir comment vos activités quotidiennes influencent l'efficacité de votre nuit ? Ou vous voulez tout simplement de nouvelles sonneries de réveil pour vous aider à mieux attaquer la journée ? Sleep Better vous permet d'améliorer la qualité de votre sommeil de façon simple et captivante ! Obtenez un aperçu de votre nuit, distinguez les phases lunaires, écrivez un journal de rêves et bien plus encore grâce à votre smartphone : Enregistrez votre durée, vos cycles et votre efficacité de sommeil à l'aide de simples clics ; Profitez de toutes les fonctions de suivi, même en Mode avion, en posant tout simplement votre appareil sur votre lit (à côté de votre coussin) […] »

3.1.2. Participants

Afin d'obtenir un effectif élevé de participants, il a été choisi de réaliser des passations en ligne. Les participants ont été recrutés sur les réseaux sociaux : Facebook, Twitter et LinkedIn. Ainsi, 305 réponses ont été collectées auprès de volontaires n'ayant jamais utilisé l'application Sleep Better. L'échantillon est composé de 236 femmes avec un âge moyen de 28.19 ans (*E-T* = 8.75) et 66 hommes avec un âge moyen de 31.11 ans (*E-T* = 11.14). Trois personnes n'ont pas souhaité transmettre leurs données démographiques.

3.1.3. Mesures

Les variables présentées précédemment peuvent être distinguées en deux catégories: les variables reliées aux qualités des produits (e.g., Utilité Perçue ; Venkatesh et al., 2003) et les qualités liées aux facteurs affectivo-motivationnels davantage rattachés aux utilisateurs comme l'Image de Soi (e.g., Kwon & Chon, 2009). Ainsi, trois échelles ont été développées pour mesurer 1) l'UX anticipée (i.e., les variables attachées aux qualités des produits, 2) les facteurs affectivo-motivationnels et 3) l'acceptabilité (voir Annexe 4).

La première échelle (échelle UX) vise à évaluer l'UX anticipée à travers quatre dimensions divisées en deux sous-dimensions. La première sous-dimension, les qualités fonctionnelles, est composée de l'Utilisabilité Perçue (4 items – e.g., L'utilisation de cette application est simple) et de l'Utilité Perçue (4 items – e.g., Je trouve que cette application m'est utile). La seconde sous-dimension, les qualités non fonctionnelles, est composée de la Stimulation Perçue (4 items – e.g., Cette

application me semble novatrice) et de la Confiance Perçue (3 items – e.g., J'ai confiance dans les informations que cette application pourrait fournir).

La seconde échelle (échelle A-M) vise à évaluer les facteurs affectivo-motivationnels à travers trois dimensions : la Motivation Intrinsèque (5 items – e.g., Indiquez dans quelle mesure chacun des énoncés suivants correspond à l'une des raisons pour lesquelles vous pourriez utiliser cette application : Parce que l'utilisation de cette application semble très fun), l'Image de Soi (4 items – e.g., Les gens pourraient m'envier d'utiliser cette application) ainsi que l'Influence Sociale (3 items – e.g., Les personnes dont j'apprécie l'opinion pourraient m'inciter à utiliser cette application).

La troisième échelle évalue les manifestations observables de l'acceptabilité à travers l'intention d'usage (4 items – e.g., La probabilité que j'utilise cette application est élevée).

Une partie des items a été construite spécialement pour ces études et d'autres ont été inspirés par plusieurs questionnaires comme l'UTAUT2 (Venkatesh et al., 2012a) ou l'Attrackdiff (Hassenzahl et al., 2010) : voir Annexe 2. La formulation de chaque item a été conçue afin d'évaluer le jugement envers un produit technologique avant usage. Pour chacun des items, les participants devaient se positionner sur une échelle de Likert en 11 points (de 0 à 10). La position de chaque item dans le questionnaire a été rendue aléatoire pour limiter les effets d'ordre de présentation des items (Benton & Daly, 1991).

3.1.4. Procédure

Les passations ont eu lieu en ligne par l'intermédiaire d'un site web spécialement développé. Tout d'abord, sur la première page, il était indiqué que nous recherchions des personnes acceptant de donner leur avis sur une application mobile, qu'ils utilisent ou non des applications mobiles. S'ils souhaitaient participer, ils étaient redirigés vers la présentation de l'application comprenant à la fois la description textuelle et les captures d'écrans (voir 3.1.1. Matériel pour plus de détails) avec l'instruction suivante : « Nous voudrions collecter votre opinion sur l'application Sleep Better. Afin de vous présenter cette application et ses caractéristiques, vous trouverez ci-dessous une description rapide ». Après avoir lu cette description, les participants devaient compléter les échelles UX et A-M ainsi que les mesures d'acceptabilité. Les

items (échelle de Likert en 11 points) des échelles étaient présentés de manière aléatoire (voir 3.1.3. Mesures pour une présentation détaillée). Enfin, les passations duraient en moyenne 7 minutes.

3.2. Résultats

Le logiciel R (R Core Team, 2016) ainsi que les librairies psych (Revelle, 2015) et lavaan (Rosseel, 2012a) ont été utilisées pour réaliser les analyses statistiques.

3.2.1. Analyses psychométriques

Pour s'assurer des qualités psychométriques des échelles de mesures après usage, des analyses statistiques ont été conduites et sont présentées ci-dessous.

Statistiques descriptives

Les statistiques descriptives ont été calculées par dimension et sont présentées dans les Tableau 3,

Tableau 4 et Tableau 5[15]. En résumé, les moyennes des dimensions varient entre 3.08 et 7.15. L'ensemble des statistiques descriptives par items est reporté en Annexe 9.

Tableau 3. Statistiques descriptives de l'échelle UX sur l'échantillon Sleep Better

Dimension	Utilité	Utilisabilité	Stimulation	Confiance
Moyenne	5.38	7.15	6.40	4.85
E.-T.	2.99	1.95	2.38	2.33

Tableau 4. Statistiques descriptives de l'échelle A-M sur l'échantillon Sleep Better

Dimension	Motivation Intrinsèque	Image de Soi	Influence Sociale
Moyenne	5.10	3.08	3.97
E.-T.	2.67	2.43	2.83

[15] Ces moyennes ont été calculées à des fins descriptives et ne seront pas utilisées dans les futures analyses où les dimensions évaluées sont représentées par des variables latentes.

Tableau 5. Statistiques descriptives Intention d'usage sur l'échantillon Sleep Better

Dimension	Intention d'Usage
Moyenne	4.84
E.-T.	3.13

Analyses Factorielles Exploratoires (AFE)

Tout d'abord, des Analyses Factorielles Exploratoires (EFA) ont été réalisées. La détermination du nombre de facteurs à extraire est basée sur les valeurs propres (*eigenvalue*). Au regard de la littérature, des corrélations sont attendues entre les facteurs mesurés (e.g., Ma & Liu, 2004). Par conséquent une rotation oblique Promax (Abdi, 2003) a été utilisée. Ainsi, une AFE par échelle a été réalisée et les résultats sont présentés en Annexe 11. Ces analyses montrent que les items saturent uniquement sur les dimensions préalablement attendues. En effet, les items saturent au-dessus de .30 uniquement sur les dimensions qu'ils sont censés mesurer.

Analyses Factorielles Confirmatoires (AFC)

Des analyses factorielles confirmatoires (AFC) ont été conduites à la suite des AFE. Pour chacune des échelles, plusieurs modèles ont été testés et comparés : du modèle le plus restrictif (i.e., tous les items sont reliés à un facteur général) au modèle le moins restrictif (i.e., tous les items sont associés aux facteurs théoriques prédéfinis). La sélection des modèles repose sur la comparaison de modèles avec un test de comparaison du χ^2 (*Satorra-Bentler Scaled Chi-Square Difference Test*) et

l'ajustement des modèles (voir Tableau 6). La non normalité multivariée des données[16] ainsi que les données manquantes ont été prises en compte[17].

En ce qui concerne l'échelle UX, les modèles suivants ont été comparés :

- M0 : modèle à 1 facteur où tous les items saturent un facteur général.
- M1 : modèle à 2 facteurs corrélés. Un premier facteur correspond aux qualités fonctionnelles (items théoriquement rattachés à l'utilité et l'utilisabilité) et un second facteur correspond aux qualités non fonctionnelles (items théoriquement reliés à la stimulation et la confiance).
- M2 : modèle à 4 facteurs corrélés. Les items sont liés à leurs facteurs théoriques prédéfinis : Utilité, Utilisabilité, Stimulation et Confiance.

La comparaison des modèles et leurs ajustements respectifs (voir Tableau 6) indiquent que le modèle avec deux facteurs corrélés (M1) est préférable au modèle à un facteur (M0) : $\Delta\chi^2$ = 207.74, Δddl = 1, p < .001. De plus, un modèle à 4 facteurs corrélés (M2) est significativement meilleur qu'un modèle à deux facteurs (M1) : $\Delta\chi^2$ = 341.05, Δddl = 5, p < .001). En conclusion, la structure en quatre facteurs corrélés apparait comme compatible avec nos données et est présentée en Figure 23.

[16] Une normalité multivariée est requise pour conduire des AFC (Anderson & Gerbing, 1988). Les résultats du test de Mardia (Mardia, 1970) indiquent une violation des hypothèses de normalité multivariée pour les deux échelles : échelle Affectivo-Motivationnelle – *Kurtosis* = 28.87, p < .001 et *Skew* = 1263.59, p < .001 ; échelle UX - *Kurtosis* = 40.18, p < .001 et *Skew* = 2847.18, p < .001. Par conséquent, un estimateur robuste (Huber-White) et des tests statistiques ajustés (Yuan-Bentler) ont été utilisés pour tenir compte de la non normalité multivariée des données (Beaujean, 2014).

[17] Une estimation par maximum de vraisemblance à informations complètes (*Full Information Maximum Likelihood*) a été employée (Enders & Bandalos, 2001) pour traiter les données manquantes.

Tableau 6. Comparaison des modèles d'AFC pour l'échelle UX[18]

Modèle	#par	χ^2	ddl	$\frac{\chi^2}{ddl}$	CFI	TLI	RMSEA	SRMR	BIC
M0	48	1497.86	104	14.41	.536	.465	.210	.169	19860
M1	49	1138.01	103	11.25	.656	.599	.182	.165	19305
M2	54	133.26	98	1.36	.988	.986	.034	.031	17795

Comparaison de modèles	$\Delta\chi^2$	Δddl	p
M0 vs. M1	207.74	1	< .001
M1 vs. M2	341.05	5	< .001

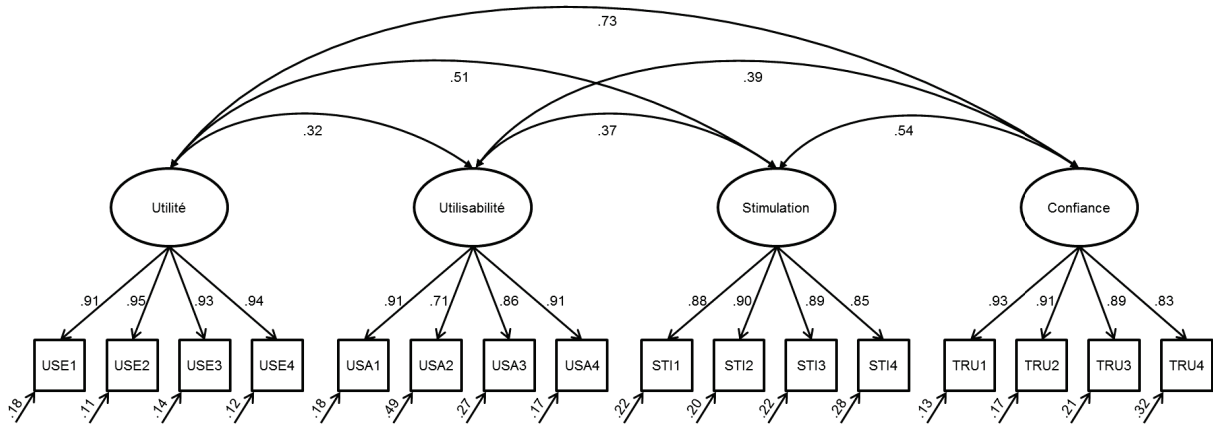


Figure 23. Résultats AFC pour l'échelle UX[19]

En ce qui concerne l'échelle A-M, les modèles suivants ont été comparés :

- M0 : modèle à 1 facteur où tous les items saturent un facteur général.
- M1 : modèle à 3 facteurs corrélés. Les items sont rattachés à leurs facteurs théoriques prédéfinis : Motivation, Image de Soi et Influence Sociale.

La comparaison des modèles et l'ajustement des modèles (voir Tableau 7) indiquent qu'un modèle avec trois facteurs (M1) est préférable au modèle à un facteur

[18] Les différences de χ^2 sont calculées sur les estimations originelles de χ^2 et non sur les estimations robustes.

[19] Tous les liens et corrélations sont significatives à $p < .001$.

(M0) : $\Delta\chi^2$ = 482.77, Δddl = 3, p < .001. Cependant, l'ajustement du modèle M1 n'est pas satisfaisant. Une analyse des covariances résiduelles montre une covariance importante entre l'item IM1 (« Indiquez dans quelle mesure chacun des énoncés suivants correspond à l'une des raisons pour lesquelles vous pourriez utiliser cette application : parce que l'utilisation de cette application semble très fun ») et l'item IM2 (« Indiquez dans quelle mesure chacun des énoncés suivants correspond à l'une des raisons pour lesquelles vous pourriez utiliser cette application : Parce qu'utiliser cette application pourrait m'amuser beaucoup »). L'ajout de cette covariance à la modélisation améliore significativement le modèle ($\Delta\chi^2$ = 35.44, Δddl = 1, p < .001) et permet d'obtenir un ajustement satisfaisant (voir Tableau 7). Cette covariance résiduelle peut s'expliquer par une formulation très proche entre les items (Hamm & Hoffman, 2016). En conclusion, la structure en trois facteurs corrélés semble se retrouver dans nos données et est présentée en Figure 24.

Tableau 7. Comparaison des modèles d'AFC pour l'échelle A-M[20]

Modèle	#par	χ^2	ddl	$\frac{\chi^2}{ddl}$	CFI	TLI	RMSEA	SRMR	BIC
M0	36	732.25	54	13.56	.598	.508	.203	.118	15865
M1	39	195.17	51	3.83	.914	.889	.096	.046	15108
M2	40	82.10	50	1.45	.981	.975	.046	.035	14947

Comparaison de modèles	$\Delta\chi^2$	Δddl	p
M0 *vs*. M1	485.85	3	< .001
M1 *vs*. M2	35.62	1	< .001

[20] Les différences de χ^2 sont calculées sur les estimations de χ^2 originales et non sur les estimations robustes.

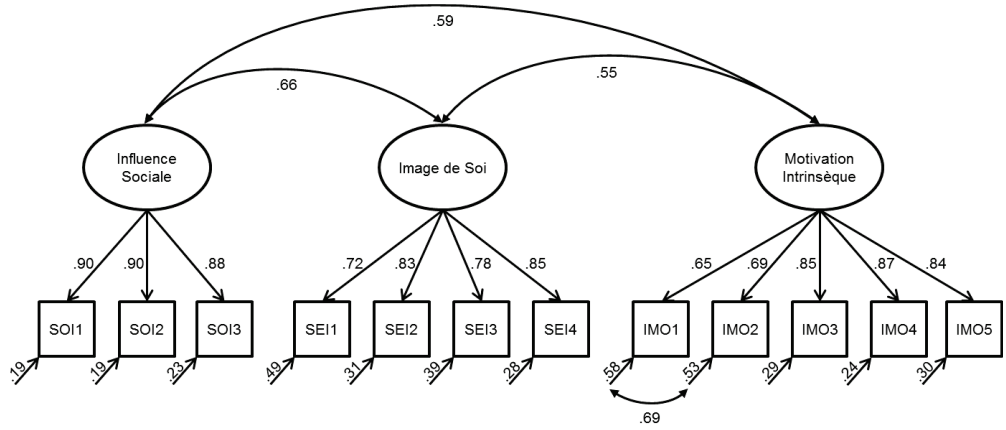


Figure 24. Résultats AFC pour l'échelle A-M[21]

Corrélations entre facteurs et alphas de Cronbach

Sur la base des modèles sélectionnés ci-dessus (modèle à 4 facteurs pour l'échelle UX et modèle à 3 facteurs pour l'échelle A-M), les corrélations entre les facteurs ainsi que les alphas de Cronbach ont été estimés. Les résultats montrent une excellente consistance interne pour tous les items (*alpha de Cronbach* ≥ .87) : voir Tableau 8.

Tableau 8. Fiabilité et corrélations entre les variables latentes[22]

	1	2	3	4	5	6	7	8
1. Utilité	(.96)							
1. Utilisabilité	.33	(.91)						
2. Stimulation	.51	.37	(.93)					
3. Confiance	.73	.39	.54	(.94)				
4. Influence Sociale	.62	.23	.36	.51	(.92)			
5. Image de Soi	.54	.28	.38	.48	.66	(.87)		
6. Motivation	.81	.34	.57	.64	.59	.54	(.90)	
7. Intention d'Usage	.93	.35	.53	.71	.61	.53	.81	(.96)

[21] Tous les liens et corrélations sont significatifs à $p < .001$.

[22] Toutes les corrélations sont significatives à $p < .001$. Les alphas de Cronbach sont entre parenthèses.

3.2.2. Modèles à équations structurelles

L'effet de chaque variable évaluée sur l'intention d'usage a été testé à l'aide des modèles à équations structurelles (SEM). Cette méthode permet d'estimer simultanément les liens supposés entre les variables et d'évaluer l'ajustement du modèle aux données (se référer au handbook de Hoyle (2012) pour une explication détaillée).

Au sein du premier modèle testé, il est supposé que toutes les variables influencent l'intention d'usage. Ce premier modèle (M1) possède un ajustement satisfaisant (*#par* = 122, χ^2 = 606.09, *df* = 405, $\frac{\chi^2}{dll}$ = 1.50, *CFI* = .972, *TLI* = .968, *RMSEA* = .040, *SRMR* = .038, *BIC* = 35555). Cependant, plusieurs liens apparaissent comme non significatifs. Afin de construire le modèle le plus parcimonieux (i.e., le modèle le plus simple) (Raykov & Marcoulides, 1999), d'autres modèles ont été estimés en supprimant un par un les liens non significatifs, tout en s'assurant d'une non dégradation significative du modèle. Ainsi, la suppression des liens non significatifs conduit à un modèle (*#par* = 177, χ^2 = 610.39, *df* = 410, $\frac{\chi^2}{dll}$ = 1.48, *CFI* = .972, *TLI* = .969, *RMSEA* = .040, *SRMR* = .038, *BIC* = 35530) où seules la Motivation Intrinsèque (β = .21) et l'Utilité Perçue (β = .84) auraient une influence positive sur l'intention d'usage (89 % de variance expliquée) : voir Figure 25.

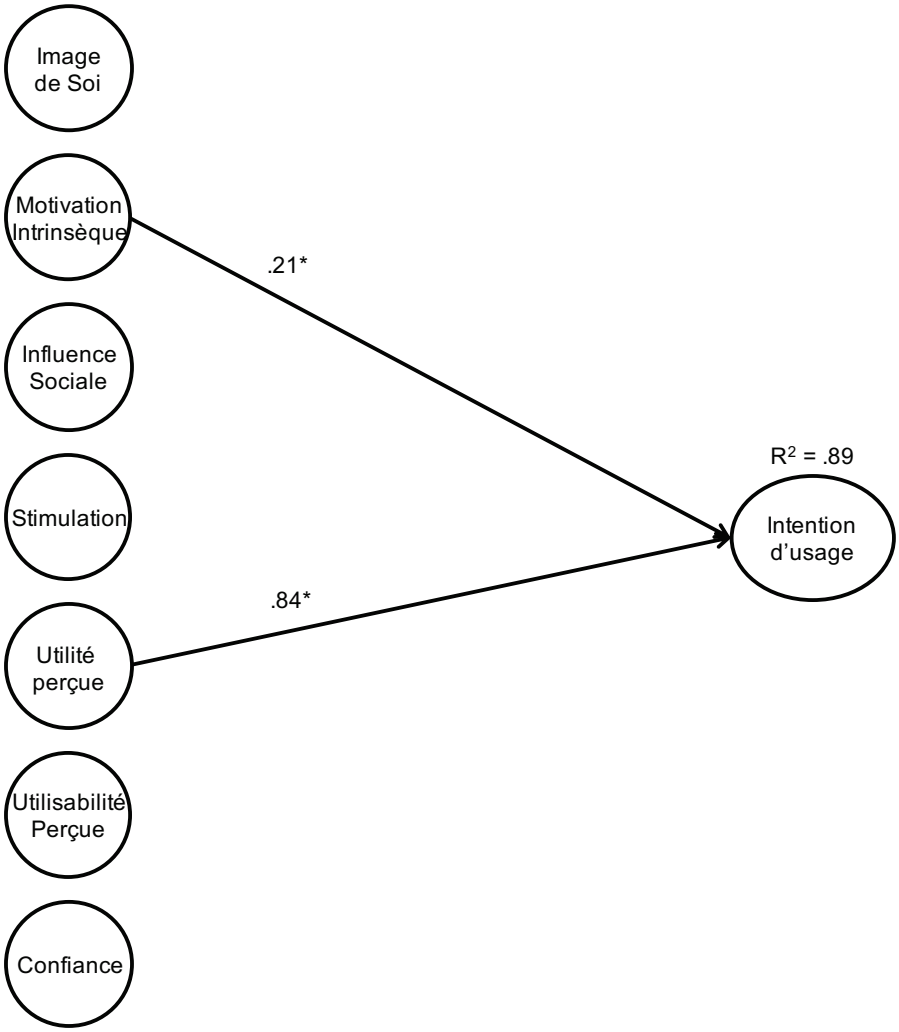


Figure 25. Modèle avec les liens estimés (modèle le plus parcimonieux) dans l'échantillon Sleep Better

3.3. Conclusion

Dans la présente étude, les variables rattachées à l'acceptation et à l'UX ont été utilisées pour modéliser le jugement avant usage. Tout d'abord, les qualités psychométriques des échelles ont été vérifiées. Les analyses montrent des propriétés satisfaisantes et par conséquent les données récoltées ont pu être utilisées pour modéliser le jugement utilisateur. Concernant les variables reliées à la perception des qualités du produit, seule l'Utilité Perçue semble avoir un effet significatif. À côté de ces qualités de produits perçues, l'influence de plusieurs variables affectivo-motivationnelles (Image de Soi, Influence Sociale et Motivation Intrinsèque) était attendue. Cependant, seule la Motivation Intrinsèque a un effet significatif sur l'intention d'usage.

En résumé, seuls des effets de l'Utilité Perçue et de la Motivation Intrinsèque sur l'intention d'usage ont été mis en avant au sein de cette première étude. Ces résultats diffèrent de ceux établis dans le contexte du jugement après usage, notamment dans l'UTAUT2 (Venkatesh et al., 2012a) et le modèle d'Hassenzahl (2005a). Ainsi, une nouvelle étude a été conduite afin de s'assurer de ces premiers résultats ainsi que de la stabilité de modèles sur une autre application mobile.

4. Étude 2 : Application Google Fit

4.1. Méthode

4.1.1. Matériel

Les participants ont évalué l'application Google Fit, une application mobile grand public de fitness *Google Fit*[23]. Une présentation courte de l'application a été proposée aux participants de manière similaire à l'étude 1. Elle est composée d'un paragraphe textuel décrivant les fonctions de l'application ainsi que plusieurs captures d'écran de l'application : voir ci-dessous (présentation complète en Annexe 8).

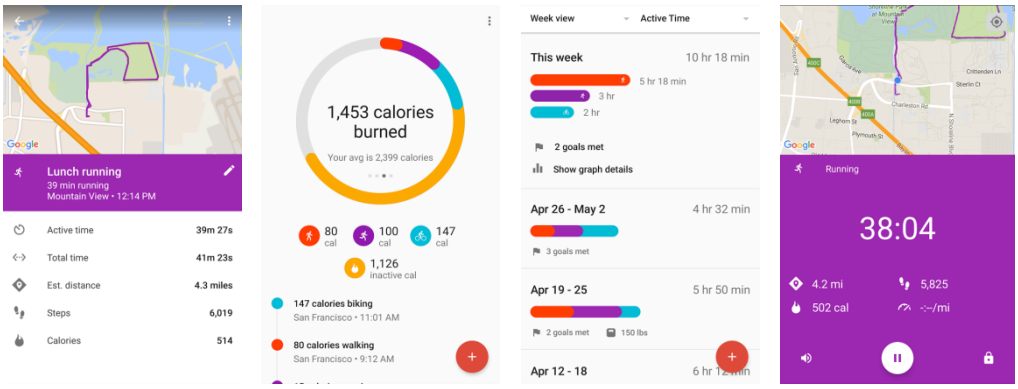


Figure 26. Captures d'écran de l'application Google Fit

« Effectuez le suivi de votre activité en toute simplicité. Lorsque vous marchez, courez ou pédalez dans la journée, votre téléphone ou votre montre Android Wear enregistrent automatiquement les données associées à votre activité avec Google Fit. ; Obtenez des informations immédiates. Découvrez en temps réel les statistiques liées à vos sessions de course, à vos promenades et à vos sorties à vélo. Google Fit

[23] Application développée par Google disponible sous Android : https://fit.google.com/.

enregistre votre vitesse, votre rythme, l'itinéraire suivi, l'altitude et d'autres informations pour vous aider à rester motivé et à atteindre vos objectifs […] ».

4.1.1. Participants

A l'instar de l'étude sur l'application mobile Sleep Better, les participants ont été recrutés sur les réseaux sociaux Facebook, Twitter et LinkedIn. Ainsi, 247 réponses ont été collectées auprès de volontaires n'ayant jamais utilisé l'application *Google Fit*. L'échantillon est composé de 172 femmes avec un âge moyen de 25.40 (*E-T* = 6.54) et 68 hommes avec un âge moyen de 29.87 (*E-T* = 9.98). 7 participants n'ont pas fourni leurs données démographiques.

4.1.2. Mesures et protocole

Les mêmes outils de mesure ainsi que le même protocole que pour l'étude 1 ont été utilisés (description textuelle et images – voir Annexe 8).

4.2. Résultats

4.2.1. Statistiques descriptives

Les statistiques descriptives ont été calculées par dimension[24] et sont présentées dans les Tableau 9, Tableau 10 et

Tableau 11. En résumé, les moyennes des dimensions varient entre 3.32 et 7.24. L'ensemble des statistiques descriptives par items est reporté en Annexe 10.

Tableau 9. Statistiques descriptives de l'échelle UX sur l'échantillon Google Fit

Dimension	Utilité	Utilisabilité	Stimulation	Confiance
Moyenne	5.10	7.24	4.02	5.65
E.-T.	2.65	1.90	2.31	2.21

[24] Ces moyennes ont été calculées à des fins descriptives et ne seront pas utilisées dans les futures analyses où les dimensions évaluées sont représentées par des variables latentes.

Tableau 10. Statistiques descriptives de l'échelle A-M sur l'échantillon Google Fit

Dimension	Motivation Intrinsèque	Image de Soi	Influence Sociale
Moyenne	4.12	3.32	3.79
E.-T.	2.69	2.62	2.82

Tableau 11. Statistiques descriptives de l'Intention d'usage sur l'échantillon Google Fit

Dimension	Intention d'Usage
Moyenne	4.14
E.-T.	2.81

4.2.2. Analyses psychométriques : invariances factorielles

Afin de s'assurer de la stabilité factorielle au travers des produits technologiques, les échelles ont été testées au sein d'un second échantillon. Pour cela, de nouveaux participants ont évalué une autre application, Google fit. Ensuite, les modèles s'ajustant aux données collectées sur l'application Sleep Better ont été testés sur ce nouvel échantillon. Enfin, l'invariance des paramètres (Milfont & Fischer, 2015) a été analysée pour s'assurer de la stabilité des paramètres au travers des échantillons[25].

[25] Du point de vue des SEMs, l'évaluation de l'invariance consiste à imposer une égalité entre certains paramètres au travers des échantillons lors de l'estimation du modèle. L'évaluation de l'invariance vise à s'assurer que le même score est mesuré au travers des échantillons pour un même score vrai (Meredith, 1993 ; A. D. Wu, Li, & Zumbo, 2007). Appliqué à l'analyse factorielle, cela signifie que les paramètres du modèle (e.g., les saturations entre les variables latentes et les items) doivent être similaires au travers des échantillons (A. D. Wu et al., 2007). Quatre niveaux d'invariances sont généralement considérés (Meredith, 1993) : invariance de configuration, invariance faible, invariance forte et invariance stricte. L'invariance de configuration revient à contraindre le modèle structural à être identique dans tous les échantillons. L'invariance faible consiste à contraindre à l'égalité sur les saturations au travers des échantillons. L'invariance forte consiste à contraindre les saturations et les intercepts à l'égalité au travers des

Pour l'échelle UX, le modèle en 4 facteurs corrélés (i.e., Utilité, Utilisabilité, Stimulation, Confiance) indique un ajustement satisfaisant : $\frac{\chi^2}{ddl}$ = 1.50, *CFI* = .979, *TLI* = .975, *RMSEA* = .045, *SRMR* = .032 et *BIC* = 14376. Sur la base de ce modèle, l'invariance factorielle a été étudiée au travers de l'évaluation de plusieurs modèles. Premièrement, l'invariance de configuration (modèle M0) et l'invariance faible (modèle M1) sont confirmées dans les analyses. Cependant, la comparaison apparait comme significative entre les modèles correspondant à une invariance factorielle faible (modèle M1) et une invariance factorielle forte (modèle M2a). Cela indique qu'un ou plusieurs intercepts sont non invariants. Par conséquent, les indices de modification ont été explorés afin de déterminer les intercepts à libérer dans les estimations. Sur cette base, plusieurs modèles ont été testés en libérant pas à pas ces intercepts. Ainsi, le modèle M2b apparait comme invariant lorsqu'une invariance partielle est imposée pour les intercepts de quatre items (quest16, quest15, quest8, quest14). Enfin, la comparaison suivante (M2b vs M3) indique une invariance stricte (modèle M3) entre les échantillons lorsque les intercepts de ces 3 items sont estimés.

échantillons. Enfin, l'invariance stricte consiste à contraindre les saturations, les intercepts et les variances résiduelles à être identiques entre les échantillons.

Tableau 12. Évaluation de l'invariance pour l'échelle UX[26]

Modèle	#par	χ^2	df	$\Delta\chi^2$	Δdf	p	CFI	RMSEA	SRMR	BIC
M0	108	252.15	196				.990	.033	.033	31752
M1	96	269.74	208	17.90	12	.12	.989	.033	.039	31699
M2a	84	308.15	220	51.05	12	<.001*	.985	.038	.044	31670
M2b	88	279.81	216	9.89	8	.27	.989	.033	.040	31657
M3	72	293.73	232	17.64	16	.35	.989	.031	.039	31618

La même approche a été utilisée pour l'échelle A-M. Ainsi, le modèle sélectionné lors des analyses sur l'échantillon Sleep Better a été testé sur ce nouvel échantillon. Les résultats indiquent un ajustement satisfaisant : $\frac{\chi^2}{df}$ = 1.50, *CFI* = .963, *TLI* = .951, *RMSEA* = .066, *SRMR* = .048, *BIC* = 12505. Par la suite, l'invariance de mesure a été évaluée (l'ensemble des comparaisons de modèles est présenté dans le Tableau 13). Tout d'abord une invariance de configuration (modèle M0) est retrouvée dans les données. Cependant, une différence significative apparait entre le modèle d'invariance de configuration (modèle M0) et le modèle d'invariance faible (modèle M1a) indiquant la présence de saturations instables entre les échantillons. Par conséquence, les indices de modification ont été examinés afin de détecter quelles saturations sont à estimer librement. Ainsi, le modèle M1b a été estimé en imposant une invariance partielle sur la saturation d'un item (quest31) et ne montre pas de différence significative avec le modèle d'invariance de configuration. Cependant, une différence significative apparait entre le modèle d'invariance faible (modèle M1b) et le modèle d'invariance forte (modèle M2), ce qui indique qu'un ou plusieurs intercepts d'items sont non invariants. Cela est confirmé par l'analyse des indices de modification. Ainsi, un modèle avec une invariance partielle pour 3 intercepts d'items a été trouvée (quest34, quest23, quest32) en estimant de nouveaux modèles en libérant pas à pas les intercepts. Enfin, en comparant ce modèle au modèle d'invariance stricte (modèle

[26] M0 : invariance de configuration ; M1 : invariance faible ; M2a : invariance forte ; M2b : invariance forte avec une invariance partielle pour les intercepts de 4 items ; M3 : invariance stricte.

M3), la différence apparait comme non significative, ce qui indique une invariance stricte en dehors des invariances partielles mentionnées préalablement.

Tableau 13. Comparaison des modèles d'invariances factorielles pour l'échelle A-M[27]

Modèle	#par	χ^2	df	$\Delta\chi^2$	Δdf	p	CFI	RMSEA	SRMR	BIC
M0	80	167.08	100				.984	.050	.041	27134
M1a	71	187.32	109	23.23	9	.005*	.981	.051	.050	27098
M1b	72	180.27	108	13.097	8	.11	.983	.050	.045	27095
M2	63	236.39	117	104.489	9	<.001	.972	.061	.052	27101
M2b	66	189.06	114	8.28	6	.218	.982	.049	.045	27065
M3	54	209.69	126	20.45	12	.059	.980	.050	.048	27028

En conclusion, en dehors de quelques paramètres non invariants, les échelles UX et A-M ont des qualités psychométriques stables au travers des échantillons et tout à fait satisfaisantes. Ces résultats permettent de valider l'utilisation de ces questionnaires. Ainsi, les données qui sont récoltées au travers de ces échelles seront utilisées pour d'autres analyses comme la modélisation des liens entre les variables.

4.2.3. Modélisations statistiques

L'intention d'usage avant usage a été modélisée en utilisant la même approche que pour l'étude 1. Le modèle complet (i.e., toutes les variables évaluées influencent l'intention comportementale) indique un ajustement satisfaisant (*#par* = 122, χ^2 = 612.36, *df* = 405, $\frac{\chi^2}{dll}$ = 1.51, *CFI* = .963, *TLI* = .958, *RMSEA* = .046, *SRMR* = .049, *BIC* = 29449). Cependant, plusieurs liens apparaissent comme non significatifs. En supprimant une par une les relations non significatives tout en s'assurant de la non dégradation de l'ajustement, seules la Motivation Intrinsèque et l'Utilité Perçue ont un effet positif sur l'intention d'usage avec un ajustement

[27] M0 : invariance de configuration ; M1a : invariance faible ; M1b : invariance faible avec une invariance partielle pour une saturation ; M2 : invariance forte ; M2b : invariance forte avec une invariance partielle pour 3 intercepts ; M3 : invariance stricte.

satisfaisant (*#par* = 117, χ^2 = 619.88, *df* = 410, $\frac{\chi^2}{dll}$ = 1.51, *CFI* = .963, *TLI* = .958, *RMSEA* = .046, *SRMR* = .049, *BIC* = 29430).

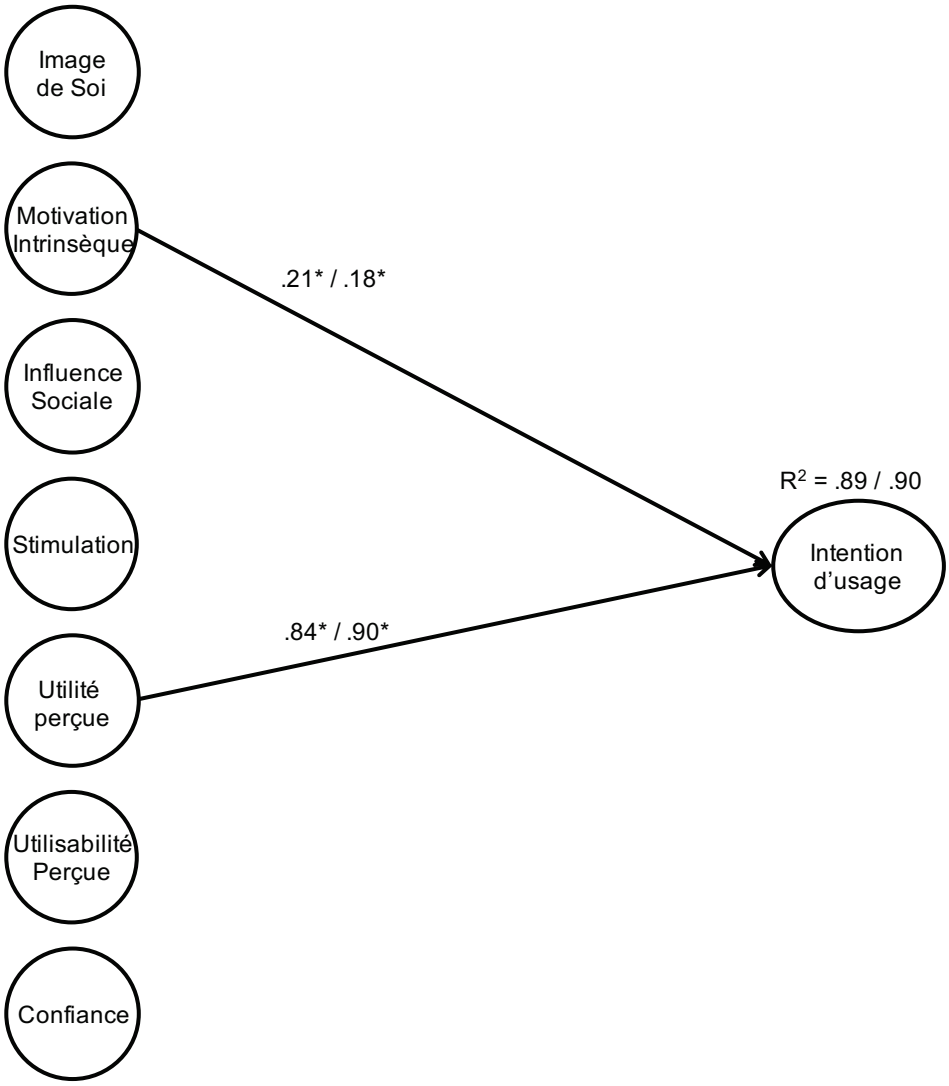


Figure 27. Modèle avec les liens estimés (modèle le plus parcimonieux) dans les deux échantillons[28]

L'étude 2 a été conduite pour s'assurer de la reproductibilité des résultats obtenus dans l'étude 1. En conclusion, les résultats sont similaires entre les deux échantillons. En effet, seules l'Utilité Perçue et la Motivation Intrinsèque auraient un

[28] Les nombres sur la gauche correspondent aux résultats sur l'échantillon Sleep Better et les nombres sur la droite correspondent aux résultats sur l'échantillon Google Fit.

effet positif sur l'intention d'usage (voir Figure 27). De plus, la variance expliquée de l'intention apparait comme élevée dans les deux échantillons avec respectivement 89 % pour l'application Sleep Better et 90 % pour l'application Google Fit.

5. Discussion

Au sein de ce chapitre, les influences des qualités perçues du produit ainsi que des facteurs affectivo-motivationnels sur l'intention comportementale avant usage ont été explorées. Pour cela, des échelles de mesure ont tout d'abord été construites. Les analyses psychométriques sur les échelles réalisées au travers de deux échantillons montrent des résultats tout à fait satisfaisants. En effet, en dehors de quelques paramètres non invariants, les indicateurs statistiques indiquent des ajustements satisfaisants. Ainsi les données collectées grâce à ces échelles ont pu servir à estimer les effets des dimensions mesurées sur l'intention d'usage.

Concernant les qualités perçues des produits, seule l'Utilité Perçue semble avoir un effet significatif positif dans les deux échantillons. Ce résultat est cohérent avec les méta-analyses sur le jugement après usage (e.g., Dwivedi et al., 2011 ; Khechine et al., 2016 ; King & He, 2006 ; Ma & Liu, 2004 ; Schepers & Wetzels, 2007). Cependant, selon les théories de l'acceptation (F. D. Davis, 1989 ; Venkatesh et al., 2003) et les résultats de ces précédentes études, l'intention d'usage devrait également être influencée par l'Utilisabilité Perçue. Or, cette variable ne semble pas adaptée dans un contexte avant usage selon les résultats obtenus dans cette étude. L'évaluation de l'utilisabilité sans interaction avec le système pourrait être difficile pour les participants et par conséquent pourrait ne pas influencer le jugement envers un produit avant usage. Des résultats similaires ont également été trouvés dans l'étude de Šumak & Šorgo (2016) sur l'intention d'usage envers des tableaux blancs interactifs de non utilisateurs. De plus, l'influence de l'Utilité Perçue est très élevée (i.e., coefficients bêta supérieurs à .80 dans les deux échantillons). Cette forte relation statistique pourrait alors absorber la majorité de la variance de l'intention d'usage.

En parallèle des qualités perçues des produits, l'effet de plusieurs variables affectivo-motivationnelles (Image de Soi, Influence Sociale et Motivation Intrinsèque) a été testé. Néanmoins, selon les résultats obtenus, seule la Motivation Intrinsèque a un effet significatif au sein des deux échantillons. L'Image de Soi et l'Influence Sociale

ne semblent pas être pertinentes dans le cadre du jugement avant usage envers une application mobile.

Plusieurs hypothèses peuvent être formulées pour expliquer ces résultats. Tout d'abord, ces variables peuvent être affectées par le biais de désirabilité sociale avec une tendance à ne pas répondre honnêtement pour défendre l'égo (Fisher, 1993) ou alors à la tendance à répondre comme socialement attendu (Mortel & Francisca, 2008). Ainsi, les participants pourraient ne pas répondre honnêtement aux items reliés à ces variables, ce qui les rendrait non pertinentes pour la modélisation statistique. De plus, des facteurs modérateurs (e.g., âge, contexte) pourraient également influencer l'effet de l'Influence Sociale. Par exemple, Venkatesh et ses collaborateurs (2003) détectent un effet significatif seulement dans un contexte d'utilisation obligatoire mais pas dans un contexte d'utilisation volontaire. Toutefois, les facteurs modérateurs ne sont pas considérés dans les méta-analyses sur le jugement après usage qui montrent un effet de l'Influence Sociale (e.g., Khechine et al., 2016). La présence ou l'absence des utilisateurs dans son entourage pourrait influencer l'effet de l'Influence Sociale.

Plus spécifiquement, par rapport à l'Influence Sociale, la présence ou l'absence de popularité de l'application pourrait expliquer l'absence d'effet. Si les participants sont uniquement entourés par des non utilisateurs de l'application, l'Influence Sociale pourrait être faible ou nulle.

Concernant l'Image de Soi, le contexte volontaire pourrait également expliquer l'absence d'effet : les précédentes études (e.g., J. D. Jackson et al., 2013 ; Kwon & Chon, 2009) montrent un effet de l'Image de Soi dans un contexte d'usage obligatoire : par exemple Jackson et collaborateurs (2013) ont étudié les facteurs influençant l'adoption d'un nouveau système auprès de personnels hospitaliers. Néanmoins, le facteur modérateur volontarisme a été supprimé de l'UTAUT2 (Venkatesh et al., 2012a) car il n'apparaissait pas comme pertinent. D'autres recherches comme celle conduite par Hsu, Lu, & Hsu (2007) sur l'intention comportementale envers l'internet mobile n'ont relevé aucun effet significatif de l'Image de Soi.

En résumé, seules les variables Utilité Perçue et Motivation Intrinsèque jouent un rôle significatif dans la formation de l'intention d'usage. Ces résultats peuvent être analysés au travers de la théorie de la motivation qui joue un rôle important dans le comportement humain (oo, Han, & Huang, 2012). Selon la théorie de Ryan & Deci (2000), la motivation serait composée de deux éléments : la motivation extrinsèque (i.e., renvoie au fait de faire quelque chose pour une raison extérieure) et la Motivation

Intrinsèque (i.e., renvoie au fait de réaliser une action parce qu'elle est intéressante en soi). Ainsi, l'Utilité Perçue pourrait être considérée comme une forme de motivation extrinsèque (R. M. Ryan & Deci, 2000 ; oo et al., 2012). Par conséquent, il peut sembler logique de retrouver, au sein de ces résultats, un effet de ces deux composantes de la motivation sur l'intention comportementale.

Enfin, l'absence ou la présence des effets supposés pourrait aussi s'expliquer par le produit évalué mais également l'échantillon. En effet, seules des applications mobiles ont été évaluées par des jeunes personnes. Ces résultats pourraient être différents avec d'autres produits (e.g., des produits moins utilisables) ou dans d'autres échantillons (e.g., personnes plus âgées).

6. Limites et futurs travaux

Ces travaux sont la première étape dans des recherches sur l'influence des qualités perçues des produits et des variables affectivo-motivationnelles sur l'intention comportementale avant usage.

Tout d'abord, au sein de ces études, le jugement a été collecté seulement avant usage et des différences apparaissent par rapport aux études précédentes sur le jugement après usage.

De plus, dans les études courantes, seul un nombre limité de facteurs a été évalué. D'autres facteurs pourraient avoir un effet selon les études précédentes sur le jugement après usage (e.g., absorption ; Ding, Huang, & Verma, 2011). Afin de limiter le temps pour les participants, il a été nécessaire de limiter le nombre de variables évaluées. Par conséquent, les modèles les plus courants et les variables rattachées ont été sélectionnés afin de limiter le nombre d'items à remplir. De plus, pour comprendre si les différences sont dues aux produits et/ou contexte, de nouvelles études avec une évaluation avant et après usage pourraient être conduites. Enfin, il est nécessaire de conduire de nouvelles recherches sur d'autres types de produits (e.g., produit d'action ; Hassenzahl, 2005a) pour s'assurer de la stabilité du modèle structurel.

7. Conclusion

Au sein de ce chapitre, le jugement avant usage a été exploré à l'inverse de la majorité des précédentes études centrées sur le jugement après usage. Selon les

résultats sur les deux échantillons, les qualités perçues des produits et les facteurs affectivo-motivationnels diffèrent selon les résultats trouvés dans le jugement après usage. Ainsi, seules l'Utilité Perçue et la Motivation Intrinsèque semblent jouer un rôle clé dans la construction du jugement avant usage au sein de nos études. Mettre l'accent sur le jugement avant usage (i.e., acceptabilité) en évaluant les perceptions des produits et les facteurs affectivo-motivationnels avant usage semble être une approche intéressante pour comprendre l'adoption des produits dans son ensemble. En effet, dans certains cas, les personnes doivent juger un produit avant de pouvoir l'utiliser.

En conclusion, mieux appréhender les facteurs conduisant à l'adoption du produit avant utilisation semble un sujet essentiel, clé pour la compréhension du jugement utilisateur qu'il semble intéressant de développer.

Chapitre 2 : Jugement après usage

1. Introduction

De très nombreux travaux ont été conduits pour comprendre le jugement utilisateur. La plupart de ces recherches a été centrée sur l'évaluation après usage (Karahanna et al., 1999 ; Šumak & Šorgo, 2016 ; S. ang et al., 2012). Dans ce contexte, le jugement après usage est défini comme l'évaluation qui est émise par un individu envers un système après au moins une utilisation de celui-ci. Comme présenté au sein de l'introduction théorique, il existe une importante variété de modèles du jugement utilisateur envers un produit technologique après usage (e.g., UTAUT ; Venkatesh et al., 2012a). Chacun de ces modèles possède ses spécificités que ce soit en termes de variables explicatives ou de variables à expliquer.

Au sein du chapitre précédent, l'évaluation avant usage a été étudiée. Or, la construction du jugement utilisateur évolue de l'évaluation avant usage jusqu'à l'adoption à long terme. Ainsi, pour compléter les premiers travaux, une étude a été conduite pour évaluer les facteurs influençant le jugement utilisateur envers un système (i.e., application mobile) après usage en se basant notamment sur les champs théoriques de l'acceptation et de l'expérience utilisateur (UX) et sera présentée dans ce chapitre.

2. Les variables influençant le jugement après usage

Au sein de l'état de l'art, de nombreux modèles théoriques et variables rattachées ont été mis en lumière dans le contexte du jugement après usage. Ainsi, il semble nécessaire de prendre en compte des variables reliées aux champs théoriques de l'acceptation, de l'UX et de la sphère affectivo-motivationnelle (voir Introduction théorique). Afin de sélectionner les variables les plus pertinentes à prendre en compte et construire les échelles de mesure les évaluant, les recherches antécédentes dans ces champs théoriques ont été analysées et sont résumées ci-dessous.

2.1. Les qualités fonctionnelles

Selon les théories de l'UX et de l'acceptation, les utilisateurs percevraient majoritairement les produits selon deux qualités fonctionnelles : l'utilité perçue (i.e., « degré selon lequel un individu croit qu'un système peut améliorer sa performance » [notre traduction] ; F. D. Davis, 1989) et l'utilisabilité (i.e., « degré selon lequel un individu croit qu'utiliser un système lui demandera peu ou pas effort » [notre traduction] ; F. D. Davis, 1989). De nombreuses études (voir méta-analyses de Dwivedi et al., 2011 ; King & He, 2006 ; Ma & Liu, 2004 ; Schepers & Wetzels, 2007 pour revue) ont été réalisées et montrent un effet important de l'utilité et de l'utilisabilité perçue sur l'intention d'usage. De plus, les travaux sur l'UX ont montré une relation entre les qualités fonctionnelles et le jugement global envers un produit. Par exemple, Hassenzahl (2004) montre une influence positive sur le jugement. Ce résultat a été confirmé par un ensemble d'études (e.g., Aranyi & van Schaik, 2015b ; Mahlke, 2007 ; Mahlke et al., 2007 ; van Schaik & Ling, 2011).

Hypothèse 1. L'utilité perçue devrait avoir une influence positive sur l'intention comportementale après usage.

Hypothèse 2. L'utilisabilité perçue devrait avoir une influence positive sur l'intention comportementale après usage.

Suite à des travaux montrant que les utilisateurs perçoivent les produits au-delà de leurs qualités fonctionnelles (Bargas-Avila & Hornbaek, 2011), des études ont été conduites pour examiner les autres facteurs qui pourraient influencer le jugement utilisateur. Sur la base d'une analyse de la littérature, seules les variables les plus utilisées seront présentées : la stimulation, l'esthétisme et la confiance.

La stimulation est rattachée au besoin de challenge et de nouveauté. Cela serait un prérequis pour le développement personnel (Hassenzahl, 2004, 2005a). Cette variable aurait un effet positif sur le jugement global (Aranyi & van Schaik, 2015a ; van Schaik & Ling, 2008).

Hypothèse 3. La stimulation devrait avoir une influence positive sur l'intention comportementale après usage.

L'esthétisme serait également une variable importante dans la formation du jugement utilisateur. Ainsi, les humains auraient tendance à juger positivement

(« *What is beautiful is good* » ; e.g., Hassenzahl, 2004) et utilisable (« *What is beautiful is usable* » ; Tractinsky, Katz, & Ikar, 2000) les produits technologiques avec de bonnes qualités esthétiques. Il a également été mis en avant un effet de l'esthétisme sur l'intention comportementale (Hall & Hanna, 2004 ; Wang, Hong, & Lou, 2010). D'autres chercheurs (e.g., Bonnardel, Piolat, & Le Bigot, 2011) ont montré une influence de la couleur, partie intégrante de l'esthétisme (e.g., Moshagen & Thielsch, 2010), sur le jugement envers un système.

Hypothèse 4. L'esthétisme devrait avoir une influence positive sur l'intention comportementale après usage.

La confiance (i.e., croyance positive concernant la fiabilité d'un système ; Fogg & Tseng, 1999) serait également un facteur important dans l'usage des technologies (Riegelsberger et al., 2005). Un effet positif de cette variable a été démontré à plusieurs reprises (Corbitt et al., 2003 ; Gefen et al., 2003 ; Jarvenpaa et al., 2000 ; Pavlou, 2003).

Hypothèse 5. La confiance envers le système devrait avoir une influence positive sur l'intention comportementale après usage.

Dans la majorité des recherches sur l'acceptabilité et l'UX, le jugement envers un produit technologique est considéré comme basé sur une évaluation rationnelle de critères de qualités (Venkatesh & Morris, 2000). Cependant, le fonctionnement humain est plus complexe et l'inclusion de variables affectivo-motivationnelles pourrait améliorer la précision des modèles.

2.2. Les facteurs affectivo-motivationnels

Au sein de la recherche en psychologie, l'influence des émotions sur le fonctionnement humain, comme le jugement ou la prise de décision (Bechara, Damasio, & Damasio, 2000 ; Pessoa, 2008), a été démontrée à de nombreuses reprises. Ainsi, des études ont été conduites sur l'influence des émotions en IHM. Aranyi et van Schaik (2015a) ont quant à eux montré un effet négatif des émotions négatives sur le jugement envers un produit. Porat et Tractinsky (2012) ont, quant à eux, mis en avant un effet positif des émotions sur les attitudes envers les produits.

Hypothèse 6. Les émotions positives devraient avoir une influence positive sur l'intention comportementale après usage.

Hypothèse 7. Les émotions négatives devraient avoir une influence négative sur l'intention comportementale après usage.

Parallèlement aux émotions, l'identification serait un prédicteur important de l'acceptabilité et de l'UX. Elle est définie comme l'image que le produit renvoie de soi (Hassenzahl, 2004). L'identification est rattachée au besoin humain d'être socialement reconnu (Hassenzahl, 2005a). Cette variable aurait un effet positif sur le jugement envers un produit (van Schaik & Ling, 2008, 2011) et l'intention d'usage (J. D. Jackson et al., 2013 ; Kwon & Chon, 2009).

Hypothèse 8. L'image de soi devrait avoir une influence positive sur l'intention comportementale après usage.

La motivation est un champ théorique important de la psychologie pour expliquer les comportements (Deci & Ryan, 1985). Dans le domaine des IHM, les recherches se sont concentrées sur la motivation intrinsèque souvent conceptualisée à travers le plaisir perçu (e.g., Balog & Pribeanu, 2010 ; Teo et al., 1999). Ainsi, la motivation intrinsèque influencerait positivement l'intention comportementale (Koufaris, 2002 ; Lee et al., 2005 ; van Schaik & Ling, 2011).

Hypothèse 9. La motivation intrinsèque devrait avoir une influence positive sur l'intention comportementale après usage.

Au-delà de la motivation, l'effet de l'absorption, comme dimension du *flow* (O'Brien & Toms, 2013), a été étudié. Cette variable est définie comme la perception des utilisateurs du temps passé et du degré de conscience à propos de ce qui se passe autour de soi lors de l'usage d'un produit (O'Brien & Toms, 2010). Cette variable est considérée comme étant reliée positivement avec le jugement global d'un produit (O'Brien & Toms, 2013) et l'intention comportementale (Ding et al., 2011).

Hypothèse 10. L'absorption devrait avoir une influence positive sur l'intention comportementale après usage.

Enfin, l'influence sociale a été ajoutée dans les recherches sur l'acceptabilité pour faire face à la vision trop rationnelle du TAM. Plusieurs méta-analyses (Dwivedi

et al., 2011 ; Schepers & Wetzels, 2007 ; Taiwo & Downe, 2013) ont montré un effet positif sur l'intention d'usage.

Hypothèse 11. L'influence sociale devrait avoir une influence positive sur l'intention comportementale après usage.

En conclusion, l'ensemble des variables présentées ci-dessus semble jouer un rôle dans jugement utilisateur et l'adoption d'un produit. Avant de construire une nouvelle échelle de mesure, les questionnaires déjà disponibles dans la littérature ont été étudiés. Cependant, il n'existe pas à notre connaissance de questionnaire permettant d'évaluer en même temps toutes les variables répertoriées ci-dessus. Afin de développer un outil à la fois complet et uniforme, de nouveaux questionnaires ont été construits en se basant sur l'Attrakdiff (Hassenzahl et al., 2010), l'Engagement Scale (O'Brien & Toms, 2013) ou encore l'UTAUT2 (Venkatesh et al., 2012a). Tous les items ont été conçus afin d'avoir la même formulation et le même format de réponse.

3. Étude : application Waze

Afin de tester les différentes hypothèses sur les variables influençant le jugement utilisateur, il a été demandé à des participants d'évaluer l'application mobile Waze.

3.1. Méthode

3.1.1. Matériel

Les participants ont évalué l'application Waze (voir Figure 28), une application de trafic et de navigation communautaire. Cette application permet de partager en temps réel l'état du trafic et des routes. Pour évaluer le jugement après usage, il est nécessaire de recruter uniquement des participants utilisant déjà l'application. L'application Waze, utilisée par plus de 65 millions d'utilisateurs (Hall-Geisler, 2016), semblait être un candidat adapté pour cette étude. De plus, elle est distribuée gratuitement sur les deux principaux magasins d'applications (i.e., iOS et Android).



Figure 28. Capture d'écran de l'application Waze

3.1.2. Participants

206 participants ayant déjà utilisé l'application Waze ont été recrutés sur le réseau social Facebook. L'échantillon est composé de 147 femmes avec un âge moyen de 26.82 (E-T = 7.37) et 55 hommes avec un âge moyen de 29.56 (E-T = 10.54). 4 participants n'ont pas souhaité fournir leurs données démographiques.

3.1.3. Mesures

Les variables présentées précédemment peuvent être distinguées en deux catégories avec d'un côté les variables rattachées aux qualités fonctionnelles et non fonctionnelles des produits (e.g., Utilité perçue ; Venkatesh et al., 2003) et de l'autre côté les facteurs affectivo-motivationnels comme l'Image de Soi (e.g., Kwon & Chon, 2009). Ainsi, trois échelles ont été développées (voir Questionnaire Après Usage) pour mesurer 1) les qualités perçues des produits, 2) les facteurs affectivo-motivationnels et 3) l'acceptabilité.

La première échelle (échelle UX) vise à évaluer les qualités perçues du produit à travers cinq dimensions : l'Utilité perçue (4 items, e.g., « je trouve que cette application m'est utile »), l'Utilisabilité perçue (4 items, e.g., « se servir de cette application est facile »), la Stimulation (4 items, e.g., « je trouve cette application novatrice »), la Confiance perçue (3 items, e.g., « j'ai confiance dans les informations

fournies par cette application ») et l'Esthétisme (4 items, e.g., « la présentation de cette application est agréable à regarder »).

La deuxième échelle (échelle A-M) vise à évaluer les facteurs affectifs et motivationnels à travers six dimensions : la Motivation Intrinsèque (5 items, e.g., « j'utilise cette application parce qu'elle m'amuse beaucoup »), l'Image de Soi (4 items, e.g., « l'utilisation de cette application renvoie une image positive de moi »), l'Influence Sociale (3 items, e.g., « les personnes de mon entourage que j'ai l'habitude d'écouter me conseillent d'utiliser cette application »), les Habitudes (4 items, e.g., « utiliser cette application est devenu automatique pour moi »), les Émotions Positives (6 items, e.g., « quand j'utilise [cette technologie], je me sens content(e), satisfait(e) »), les Émotions Négatives (6 items, e.g., « quand j'utilise [cette technologie], je me sens agacé(e), énervé(e), exaspéré(e) »).

La troisième échelle évalue les manifestations observables de l'acceptabilité à travers l'Intention d'Usage (4 items, e.g., « j'ai l'intention de continuer à utiliser cette application »).

Les items ont été construits spécialement pour évaluer le jugement après usage en s'inspirant notamment d'autres questionnaires comme l'UTAUT2 (Venkatesh et al., 2012a), l'Attrackdiff (Hassenzahl et al., 2010) ou la PANAS (Watson et al., 1988). Pour chacun des items, les participants devaient se positionner sur une échelle de Likert en 11 points (de 0 à 10). La position de chaque item dans le questionnaire a été rendue aléatoire pour limiter les effets reliés à l'ordre de présentation des items (Benton & Daly, 1991).

3.1.4. Procédure

Les passations ont eu lieu en ligne par l'intermédiaire d'un site web spécialement développé. Les participants étaient invités à prendre part à l'étude en ligne à partir d'un message diffusé sur le réseau social Facebook. Le lien mis en ligne conduisait à une page web où il était demandé aux participants d'évaluer l'application Waze avec la consigne suivante : « *Bienvenue. Dans le cadre d'un projet de thèse, nous recherchons des personnes acceptant de donner leur avis sur Waze. Si vous avez déjà utilisé cette application, votre avis nous intéresse. Pour toute question, n'hésitez pas à me contacter : nicolas.martin@b-com.com. Si vous souhaitez participer à cette étude, il vous suffit de cliquer sur la page suivante. Répondez le plus sincèrement possible aux questions. Il n'y a pas de bonne ou de mauvaise réponse.*

Merci de répondre à toutes les questions. ». Après avoir lu cette consigne, les participants devaient compléter les échelles UX et A-M ainsi que les mesures d'acceptabilité. Les items (échelle de Likert en 11 points) des échelles étaient présentés de manière aléatoire (voir 3.1.3. Mesures pour une présentation détaillée). Enfin, les passations duraient en moyenne 12 minutes.

3.2. Résultats

Pour réaliser les différentes analyses statistiques, le logiciel R (R Core Team, 2016) ainsi que les librairies psych (Revelle, 2015) et lavaan (Rosseel, 2012a) ont été utilisées.

3.2.1. Analyses psychométriques

Les qualités psychométriques des échelles ont été évaluées à l'aide d'analyses factorielles exploratoires et confirmatoires avant de modéliser le jugement utilisateur.

Statistiques descriptives

L'ensemble des statistiques descriptives des échelles A-M et UX, ainsi que des mesures d'acceptabilité est reporté en Annexe 12. La Figure 29 présente l'histogramme de la variable intention d'usage.

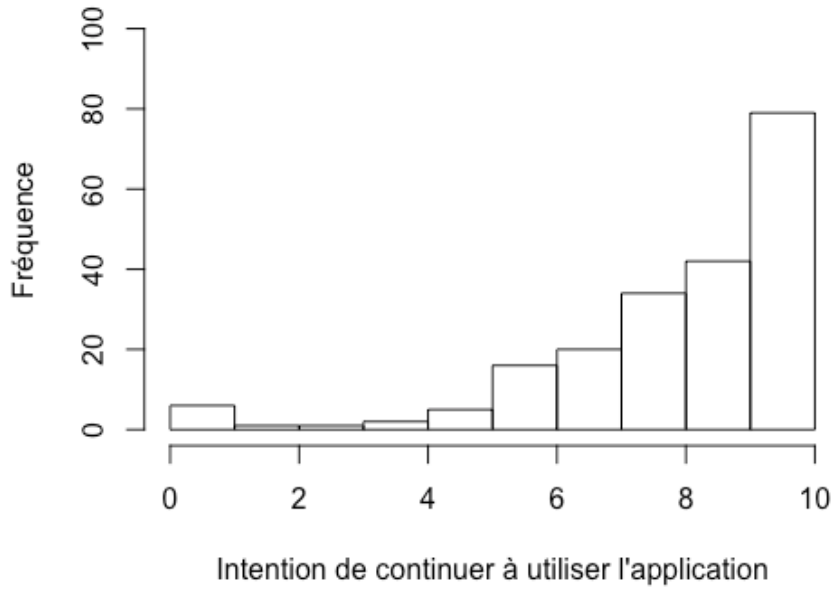


Figure 29. Histogramme des données de la variable intention comportementale

Les statistiques descriptives ont été calculées par dimension et sont présentées dans les Tableau 3,

Tableau 4 et Tableau 5[29]. En résumé, les moyennes des dimensions varient entre 3.97 et 7.15.

Tableau 14. Statistiques descriptives de l'échelle UX

Dimension	Utilité	Utilisabilité	Stimulation	Confiance	Esthétisme
Moyenne	7.80	7.90	5.97	7.81	6.35
E.-T.	2.11	2.15	2.62	1.66	2.38

Tableau 15. Statistiques descriptives de l'échelle A-M

Dimension	Motivation intrinsèque	Image de soi	Influence sociale	Absorption
Moyenne	4.12	3.32	3.79	1.79
E.-T.	2.69	2.67	2.82	2.31

Dimension	Habitudes	Emotions négatives	Emotions positives
Moyenne	4.09	1.08	4.48
E.-T.	2.68	1.95	2.67

Tableau 16. Statistiques descriptives de l'Intention d'Usage

Dimension	Intention d'usage
Moyenne	8.09
E.-T.	2.44

Analyses Factorielles Exploratoires

Des Analyses Factorielles Exploratoires (AFE) ont été réalisées sur les données récoltées avec les échelles UX et A-M après usage. Ainsi, une AFE par échelle a été

[29] Ces moyennes ont été calculées à des fins descriptives et ne seront pas utilisées dans les futures analyses où les dimensions évaluées sont représentées par des variables latentes.

réalisée[30] et les résultats sont présentés en Annexe 14. Pour l'échelle UX, l'item USE1 pour l'utilité perçue présente des saturations sur plusieurs facteurs. En ce qui concerne l'échelle A-M, les items EMP1 et EMP2 pour les émotions positives, ABS3 pour l'absorption et l'item IMO4 pour la motivation intrinsèque présentent également des saturations sur plusieurs facteurs. Par conséquent, ces items seront exclus des prochaines analyses.

Analyses Factorielles Confirmatoires

Pour chacune des échelles, plusieurs modèles d'Analyses Factorielles Confirmatoires (CFA) ont été testés et comparés : du modèle le plus restrictif (i.e., tous les items sont reliés à un facteur général) au modèle le moins restrictif (i.e., tous les items sont rattachés aux facteurs théoriques prédéfinis). La sélection des modèles repose sur deux critères : la comparaison de modèles avec un test de comparaison du χ^2 (*Satorra-Bentler Scaled Chi-Square Difference Test*) et l'ajustement des modèles (voir Tableau 17).

En ce qui concerne l'échelle UX, les modèles suivants ont été comparés :

- M0 : modèle à 1 facteur où tous les items saturent un facteur général.
- M1 : modèle à 2 facteurs corrélés. Un premier facteur correspond aux qualités fonctionnelles (items théoriquement rattachés à l'utilité perçue et l'utilisabilité perçue) et un second facteur correspond aux qualités non fonctionnelles (items théoriquement rattachés à la stimulation, la confiance et l'esthétisme).
- M2 : modèle à 5 facteurs corrélés. Les items sont rattachés à leurs facteurs théoriques prédéfinis : utilité, utilisabilité, stimulation, confiance et esthétisme.

La comparaison des modèles et leurs ajustements respectifs[31] (voir Tableau 17) indiquent que le modèle à deux facteurs corrélés (M1) est préférable ($\Delta\chi^2$ = 20.13,

[30] Le nombre de facteurs à extraire a été déterminé à l'aide d'un *scree plot* (voir Annexe 13) et une rotation oblique Promax (Abdi, 2003) a été utilisée puisque des corrélations sont attendues entre les facteurs mesurés (e.g., Ma & Liu, 2004).

[31] L'hypothèse de normalité multivariée (Anderson & Gerbing, 1988) n'est pas respectée au sein des données selon le test de Mardia (1970) : échelle UX – *Kurtosis* = 42.60, $p < .001$ et *Skew* = 3801.05, $p < .001$; échelle A-M – *Kurtosis* = 22.38, $p < .001$ et *Skew* = 8089.07, $p < .001$. Un estimateur robuste (Huber-White) et des tests statistiques

Δddl = 1, p < .001) au modèle à un facteur (M0). De plus, un modèle à cinq facteurs corrélés (M2) a un ajustement significativement meilleur ($\Delta\chi^2$ = 254.13, Δddl = 9, p < .001) qu'un modèle à deux facteurs (M1). En conclusion, la structure en cinq facteurs corrélés (voir Figure 30) apparait comme compatible avec nos données.

Tableau 17. Comparaison des modèles d'AFC pour l'échelle UX[32]

Modèle	#par	χ^2	ddl	$\frac{\chi^2}{ddl}$	CFI	TLI	RMSEA	SRMR	BIC
M0	54	736.53	135	5.46	.668	.624	.147	.109	14787
M1	55	620.69	134	4.63	.732	.693	.133	.101	14602
M2	64	134.43	125	1.08	.995	.994	.019	.035	14016

Comparaison de modèles	$\Delta\chi^2$	Δddl	p
M1 vs M2	21.02	1	< .001
M2 vs M3	284.25	9	< .001

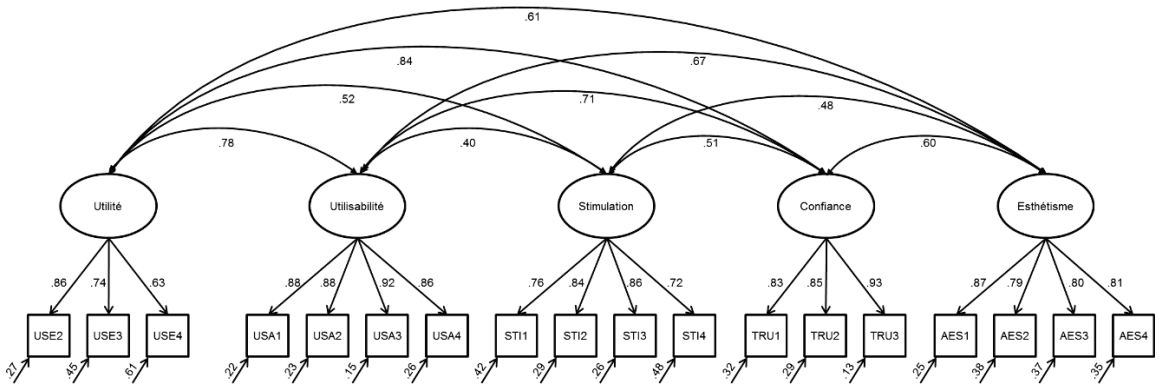


Figure 30. Résultats de l'AFC pour l'échelle UX[33]

En ce qui concerne l'échelle A-M, les modèles suivants ont été comparés :

ajustés (Yuan-Bentler) ont été utilisés (Beaujean, 2014) afin de tenir compte de la non normalité.

[32] Les différences de χ^2 sont calculées sur les estimations de χ^2 originales et non sur les estimations robustes.

[33] Tous les liens et corrélations sont significatifs à p < .001.

- M0 : modèle à 1 facteur où tous les items saturent un facteur général.
- M1 : modèle à 7 facteurs corrélés. Les items sont rattachés à leurs facteurs théoriques prédéfinis : motivation intrinsèque, image de soi, absorption, émotions négatives, émotions positives, habitudes et influence sociale.

La comparaison des modèles et l'ajustement des modèles (voir Tableau 18) indiquent que le modèle à trois facteurs (M1) est préférable ($\Delta\chi^2$ = 510.92, Δddl = 21, p < .001) au modèle à un facteur (M0). En conclusion, la structure en sept facteurs corrélés semble se retrouver dans nos données (les paramètres du modèle sont présentés en Annexe 15).

Tableau 18. Comparaison des modèles d'AFC pour l'échelle A-M[34]

Modèle	**#par**	χ^2	**ddl**	$\frac{\chi^2}{ddl}$	**CFI**	**TLI**	**RMSEA**	**SRMR**	**BIC**
M0	99	1886.77	377	5.00	.371	.323	.139	.174	27203
M1	108	478.61	356	1.34	.949	.942	.041	.062	25447

Comparaison de modèles	$\Delta\chi^2$	Δ**ddl**	**p**
M0 vs M1	510.92	21	< .001

Les indices d'ajustements ainsi que les paramètres des modèles sont relativement corrects sans être totalement satisfaisants. Cela peut notamment s'expliquer par la taille relativement faible de l'échantillon. En effet, l'échelle A-M est composée de 29 items pour 206 participants, ce qui fait un ratio relativement faible de 7 participants par item.

3.2.2. Modélisations statistiques

L'influence des variables mesurées sur l'intention comportementale a été testée à l'aide de SEMs[35]. Un premier modèle a été estimé en considérant l'ensemble des

[34] Les différences de χ^2 sont calculées sur les estimations de χ^2 originales et non sur les estimations robustes.

[35] Les SEMs permettent d'estimer simultanément les liens supposés entre les variables et d'évaluer l'ajustement du modèle aux données (se référer au Handbook de Hoyle (2012) pour une revue détaillée).

variables mesurées comme des prédicteurs de l'intention d'usage. Ce modèle offre un ajustement satisfaisant (*#par* = 122, χ^2 = 1409.88, *df* = 1146, $\frac{\chi^2}{dll}$ = 1.23, *CFI* = .937, *TLI* = .930, *RMSEA* = .033, *SRMR* = .060, *BIC* = 41474). Selon cette modélisation, seule la variable habitudes aurait une influence significative sur l'intention d'usage (β = .79, p < .001)[36]. L'influence majeure de cette variable, empruntée notamment aux travaux de Venkatesh et collaborateurs (2012a), pourrait s'expliquer par une corrélation très élevée entre l'intention comportementale et les habitudes (r = .97, p = .001). Par conséquent, il semble envisageable de considérer les habitudes plutôt comme un marqueur observable de l'adoption d'un système que comme un prédicteur de l'intention d'usage. D'autres modélisations ont ainsi été conduites en excluant la variable habitudes.

Toutes les variables mesurées en dehors de la variable habitudes ont été utilisées, dans un nouveau modèle, pour prédire l'intention comportementale. Ce modèle présente un ajustement relativement satisfaisant (*#par* = 122, χ^2 = 1157.92, *df* = 968, $\frac{\chi^2}{dll}$ = 1.20, *CFI* = .949, *TLI* = .943, *RMSEA* = .031, *SRMR* = .060, *BIC* = 39103). Cependant, selon les résultats de cette modélisation, plusieurs liens apparaissent comme non significatifs. Afin de construire le modèle le plus parcimonieux (i.e., le modèle le plus simple ; Raykov & Marcoulides, 1999), d'autres modèles ont été estimés en supprimant un par un les liens non significatifs, tout en s'assurant de la non dégradation de l'ajustement des modèles.

La comparaison des modèles successifs en supprimant les liens non significatifs conduit à un modèle où seules l'utilité perçue (β = .87, p < .001), l'utilisabilité perçue (β = .22, p = .116), l'influence sociale (β = .09, p = .055) et les émotions négatives (β = -.21, p = .030) auraient une influence sur l'intention d'usage (86 % de variance expliquée) avec un ajustement relativement satisfaisant (*#par* = 122, χ^2 = 1160.62, *df* = 975, $\frac{\chi^2}{dll}$ = 1.19, *CFI* = .950, *TLI* = .944, *RMSEA* = .030, *SRMR* = .060, *BIC* = 39069). La suppression des variables non significatives de ce modèle conduit à des dégradations significatives du modèle ou à des problèmes

[36] Coefficient bêta totalement standardisé.

d'estimation. Cela conduit à retenir ce modèle comme le plus pertinent au regard des données collectées[37] : voir Figure 31.

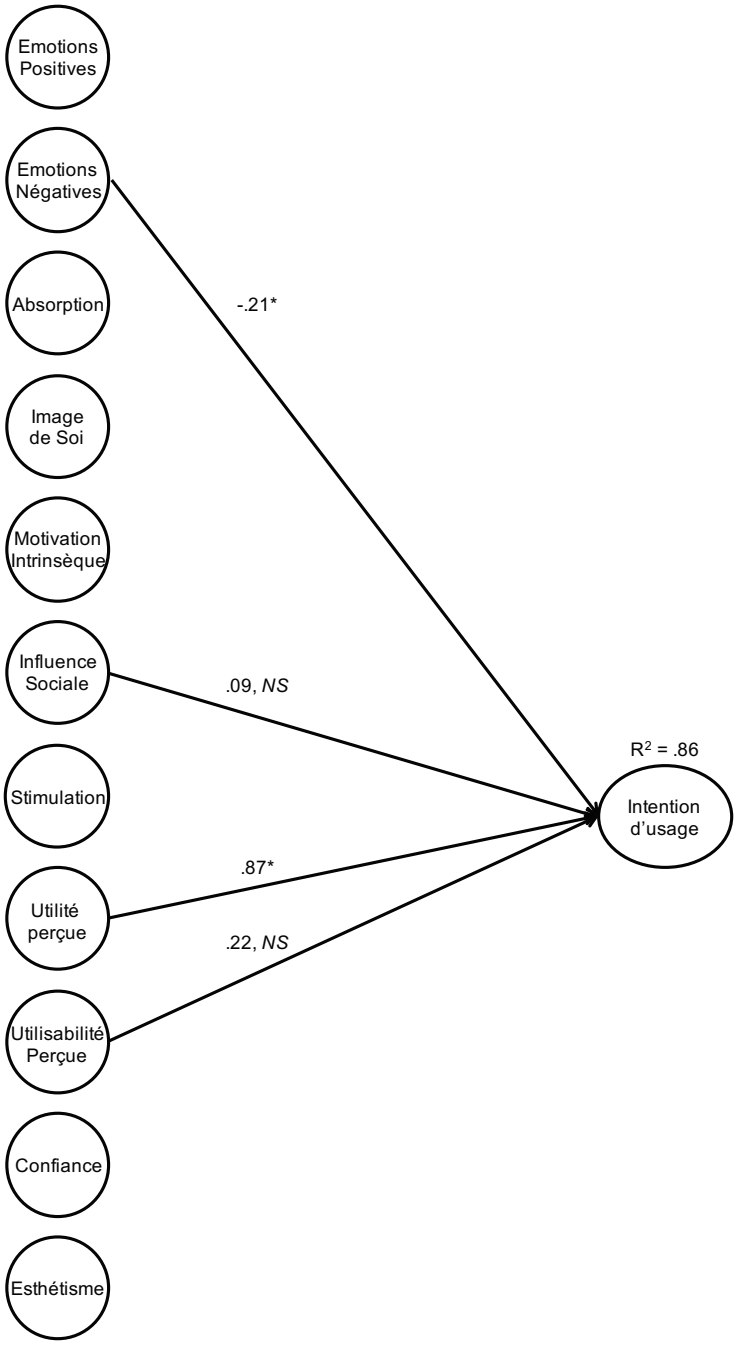


Figure 31. Modèle avec les liens estimés (modèle le plus parcimonieux)

[37] D'autres types de modélisations (modèles structuraux basés sur l'approche par les moindres carrés partiels et régressions multiples) ont été testés sur ces données et sont présentés en Annexe 16.

3.3. Discussion

Cette étude visait à discerner les variables influençant le jugement utilisateur après usage. Pour cela, il a été choisi d'intégrer les modèles d'acceptation (e.g., UTAUT ; Venkatesh et al., 2012a), d'UX (e.g., Hassenzahl, 2005a) et les facteurs affectivo-motivationnels (e.g., motivation intrinsèque ; F. D. Davis et al., 1992) dans un unique modèle afin d'obtenir une meilleure compréhension du jugement utilisateur après usage. Ainsi, l'utilité perçue, l'utilisabilité perçue, la stimulation, l'esthétisme perçu, la confiance, les émotions négatives, les émotions positives, l'influence sociale, l'image de soi, la motivation intrinsèque, l'absorption, les habitudes et l'intention d'usage ont été évalués auprès d'un échantillon de 206 utilisateurs de l'application Waze.

Dans un premier temps, les questionnaires spécialement conçus pour évaluer les dimensions présentées ci-dessus ont fait l'objet d'analyses psychométriques. Les résultats montrent des résultats encourageants mais il semble nécessaire de conduire des analyses sur de nouvelles données pour s'assurer de la reproductibilité des résultats.

Ensuite, à partir des données collectées avec ces échelles, des modélisations ont été menées en intégrant toutes les variables présentées précédemment. La variable habitudes est notamment utilisée comme variable influençant l'intention comportementale comme proposé par Venkatesh et collaborateurs (2012a). Cependant, lorsque cette dimension est incluse dans la modélisation, aucune autre variable ne présente un effet significatif sur l'intention d'usage. Cela pourrait s'expliquer par la corrélation très élevée entre l'intention d'usage et les habitudes (r = .97, p < .001). Ce résultat a amené à considérer les habitudes comme un marqueur observable de l'adoption du produit et non comme une variable explicative du comportement. Par conséquent, d'autres modèles ont été estimés en supprimant la variable habitudes. Le modèle a été de nouveau estimé en supprimant les chemins non significatifs en s'assurant de la non dégradation de l'ajustement des modèles. Cela a conduit à un modèle incluant des liens entre l'intention d'usage et l'utilité perçue, l'utilisabilité perçue, l'influence sociale ainsi que les émotions négatives. Cependant, il

est important de noter que les variables utilisabilité perçue et influence sociale n'ont pas d'effet significatif dans cette modélisation[38].

L'influence positive majeure de l'utilité perçue est cohérente avec les méta-analyses sur les modèles de l'acceptation UTAUT (Dwivedi et al., 2011 ; Khechine et al., 2016) et TAM (King & He, 2006 ; Ma & Liu, 2004 ; Schepers & Wetzels, 2007). Ce résultat est également en accord avec les travaux d'Hassenzahl et collaborateurs (Hassenzahl, 2005a ; Hassenzahl et al., 2002) sur le mode d'action. En effet, l'application évaluée dans cette expérimentation présente un but utilitaire pour ses utilisateurs. Elle vise à les aider à se rendre à une destination en limitant notamment leur temps de trajet en évitant par exemple les embouteillages. Ce mode d'usage orienté vers un objectif (« *goal-mode* » ; Hassenzahl, 2005a) aurait une influence sur le lien entre les qualités pragmatiques (e.g., utilité perçue) et le jugement global envers un produit (Hassenzahl et al., 2002). Ainsi, les qualités fonctionnelles seraient davantage corrélées au jugement global lorsque l'usage est orienté vers un but que lorsque l'usage du produit est une fin en soi (« *action-mode* » ; Hassenzahl, 2005a).

En ce qui concerne les émotions négatives, leur influence négative et significative sur l'intention d'usage est en accord avec les prédictions du modèle me-CUE (Mahlke & Minge, 2008) et les résultats de Aranyi et van Schaik (2015a).

Au sujet des variables n'ayant pas d'effet significatif sur l'intention de continuer à utiliser l'application, plusieurs hypothèses peuvent être proposées. Tout d'abord, la faible variance de la variable intention d'usage pourrait apporter une justification pour une partie des résultats. En effet, les participants souhaitent majoritairement poursuivre leur usage de l'application (médiane = 8.67 – voir Figure 29). Or, les outils pour modéliser le comportement qui ont été utilisés reposent sur l'analyse de la variance et de la covariance entre les variables. De plus, le lien très élevé entre l'utilité perçue et l'intention d'usage pourrait absorber la majorité de la variance de l'intention d'usage. Ainsi, par exemple, l'effet de l'utilisabilité perçue sur l'intention d'usage est bien étayé dans la littérature (e.g., Khechine et al., 2016 ; King & He, 2006). Néanmoins, bien que le lien soit présent dans le modèle retenu, celui-ci est non

[38] Dans les SEMs, certaines variables peuvent jouer un rôle important dans une modélisation sans pour autant avoir un effet significatif. Ce phénomène correspond à des situations où la suppression d'un lien non significatif conduit à une dégradation de l'ajustement du modèle. Ainsi, le lien non significatif est alors conservé dans le modèle.

significatif. La très large distribution de ce type d'applications (environ 65 millions d'utilisateurs pour l'application utilisée dans l'expérimentation ; Hall-Geisler, 2016) permet de supposer que les problèmes d'utilisabilité sont très réduits et donc l'application majoritairement évaluée comme utilisable. Par conséquent, la faible variance dans l'évaluation de l'utilisabilité et l'influence majeure de l'utilité offrirait un rôle réduit à l'utilisabilité dans le jugement de l'utilisateur.

De plus, une large part des travaux antécédents sur l'intention comportementale repose sur des modélisations par l'approche des moindres carrés partiels[39] (e.g., Venkatesh et al., 2003) ou des régressions linéaires multiples (e.g., F. D. Davis, 1989). Des tentatives de modélisations reposant sur ces approches ont été testées sur les données collectées durant cette étude (voir Annexe 16) et indiquent des résultats légèrement différents (e.g., effet plus faible de l'utilité perçue ; présence d'un effet significatif de l'utilisabilité). Sans remettre en question le choix de la modélisation SEM[40], ces différences de méthodologies pourraient expliquer une partie des différences entre les résultats de cette étude (e.g., pas d'effet significatif de l'utilisabilité perçue sur l'intention d'usage) et les précédents résultats de la littérature (e.g., un effet de l'utilisabilité perçue sur l'intention comportementale ; Khechine et al., 2016).

Plusieurs pistes d'améliorations sont à noter concernant cette étude. Tout d'abord, les analyses n'ont pu être conduites que sur un seul échantillon. Il semble nécessaire de poursuivre les travaux, que cela soit sur le plan des analyses psychométriques des échelles ou sur l'analyse des composantes du jugement utilisateur. En effet, les résultats pourraient être différents sur d'autres systèmes

[39] *Partial Least Square - PLS* (Tenenhaus, Vinzi, Chatelin, & Lauro, 2005).

[40] Les SEMs offrent un cadre d'analyse de données unifié (Fan, 2003) permettant notamment l'analyse du modèle de mesure (i.e., liens entre variables observables et variables non observables) et l'analyse du modèle structural (i.e., liens entre les variables non observables) (Gefen, Straub, & Boudreau, 2000). Il permet également de prendre en compte les problèmes liés aux données manquantes (Allison, 2003) ou encore à la non-normalité des données (Beaujean, 2014). Les modèles PLS (*Partial Least Square*) ou encore les régressions linéaires offrent certains avantages. Par exemple, les premiers ne font pas d'hypothèses sur les distributions de données et permettent de travailler avec des échantillons plus réduits (Tenenhaus et al., 2005). Cependant, ces méthodes ne permettent pas de tester l'ajustement du modèle aux données (Rönkkö & Evermann, 2013) ou encore de prendre en compte les erreurs de mesure.

notamment avec des qualités fonctionnelles ou non fonctionnelles plus dégradées pouvant générer de la variabilité dans l'évaluation. De plus, réaliser des études sur des systèmes pouvant être utilisés dans un autre d'usage que le mode action comme un jeu par exemple, pourrait améliorer la compréhension des liens entre l'intention d'usage et ses composantes.

4. Conclusion

Les résultats sur les données collectées après usage montrent une influence significative de l'utilité perçue et des émotions négatives ainsi qu'un effet non significatif de l'influence sociale et de l'utilisabilité perçue. Ces résultats ne sont pas totalement en accord avec les précédents travaux dans la littérature. Cela pourrait notamment s'expliquer par les données collectées. En effet, la plupart des participants souhaitaient poursuivre leur usage de l'application réduisant ainsi les variances observées sur les différentes variables mesurées. Les résultats obtenus pourraient ainsi être relativement biaisés puisque les méthodes de modélisation utilisées reposent sur la variance et la covariance dans les données.

Chapitre 3 : Évolution du jugement utilisateur au cours du temps[41]

1. Introduction

Adopter un système est un processus sur le long terme. Celui-ci débute avec les représentations initiales (i.e., représentations avant usage) et évolue au cours du temps au fur et à mesure des usages. Comprendre les facteurs qui conduisent à la poursuite de l'usage est essentiel notamment pour les entreprises qui développent des produits technologiques. En effet, un certain nombre de modèles économiques ne reposent plus uniquement sur la vente du système mais sur l'usage répété du système. L'utilisation prolongée du produit et de ses services (payants) complète les revenus obtenus lors de la vente du système. Il devient alors nécessaire que le produit satisfasse le consommateur sur le long terme. D'autres sociétés basent leurs rémunérations sur la publicité dont les bénéfices sont maximisés par l'usage prolongé et répété d'un produit (Luo, 2002). L'usage répété et prolongé peut également participer à la fidélisation globale envers la marque pouvant conduire à l'utilisation d'autres produits de celle-ci. L'intérêt pour la poursuite de l'utilisation n'est pas seulement limité aux produits commerciaux. Le principe d'engagement sur le long terme se retrouve également dans des systèmes comme les dispositifs de formation à distance dans le domaine de l'éducation (e.g., *Massive Open Online Course* - MOOC). Dans ce type de formation, il est demandé aux participants de suivre un programme sur plusieurs semaines. Or, la réussite de ces programmes dépend du suivi des participants au cours du temps. Ainsi, de nombreuses études ont été conduites afin d'identifier les facteurs facilitateurs ou inhibiteurs de l'adoption et de l'usage d'un produit technologique.

Sur le plan théorique, il est nécessaire de distinguer l'acceptabilité (i.e., jugement envers un produit avant usage ; Schuitema et al., 2010) et l'acceptation (i.e.,

[41] Ces résultats ont été publiés dans la conférence *Human-Computer Interaction International* (HCII16) à Toronto (Martin, Jamet, Erhel, & Rouxel, 2016b). De plus, une partie des données utilisées est identique à celles utilisées dans les chapitres précédents.

jugement envers un produit après usage ; Schuitema et al., 2010). Cependant, bien que cette différenciation ait été mise en avant, la majorité des travaux antérieurs a été centrée sur l'acceptation (i.e., jugement après usage). Au sein de ce chapitre, les différences entre l'évaluation avant et après usage seront explorées dans le contexte d'une application mobile. De plus, les facteurs expliquant la poursuite de l'usage d'une application mobile seront étudiés en se basant sur le cadre théorique de l'EDT (*Expectation-Disconfirmation Theory*).

1.1. Théorie des Attentes-Infirmations (EDT - *Expectation-Disconfirmation Theory*)

Le cadre théorique des attentes (*expectations*) et infirmations (*disconfirmations*) a été développé pour expliquer les comportements post-achat. À notre connaissance, l'EDT est la seule à offrir un cadre théorique pour expliquer la poursuite de l'usage (e.g., Bhattacherjee & Premkumar, 2004 ; Venkatesh, Thong, Chan, Hu, & Brown, 2011). Basée sur la théorie de la dissonance cognitive (Festinger, 1962), l'EDT décrit les processus du comportement individuel avec un produit au cours du temps. Selon cette théorie (Oliver, 1980), lorsque les clients achètent un produit, ils ont accès à des informations sur ce produit basées, par exemple, sur des affirmations des vendeurs (Bhattacherjee & Premkumar, 2004). Ces informations influenceraient leurs attentes initiales. Au fur et à mesure du temps et des usages, les utilisateurs gagnent en expériences et en informations sur le produit (Venkatesh et al., 2011). Ainsi, ils peuvent comparer leurs attentes initiales avec leurs nouvelles informations et expériences avec le produit. S'il y a des dissonances (i.e., différence(s) entre leurs connaissances et la réalité ; Bhattacherjee & Premkumar, 2004), ils peuvent être amenés à réviser leurs attitudes et comportements envers le produit (e.g., poursuivre ou non l'usage du produit). Enfin, les infirmations (i.e., différence(s) entre les attentes initiales et la réalité perçue) et les attentes initiales détermineraient la satisfaction utilisateur qui expliquerait la poursuite de l'usage (Oliver, 1980 ; Venkatesh et al., 2011).

1.2. EDT et adoption des produits technologiques

A partir du modèle de l'EDT, plusieurs études sur l'évolution du jugement envers des produits technologiques ont été conduites. Ainsi, Karahanna, Straub et Chervany (1999) ont étudié les éventuelles différences entre les participants n'ayant pas encore

testé le produit et les utilisateurs actuels. Ils constatent des différences entre ces deux groupes de participants. Par exemple, l'utilisabilité perçue serait un prédicteur significatif pour les potentiels utilisateurs mais pas pour les actuels utilisateurs. Dans une autre étude, Bhattacherjee et Premkumar (2004) ont étudié l'évolution de l'utilisabilité perçue d'un produit à trois moments d'évaluation : avant usage, trois semaines après implémentation et 10 semaines après implémentation. Ils ont montré dans deux échantillons indépendants une diminution significative de l'utilisabilité perçue entre la mesure avant usage et la mesure réalisée après 3 semaines d'usage mais pas de différence entre l'évaluation à 3 semaines et à 10 semaines. Enfin, Venkatesh et collaborateurs (Venkatesh et al., 2011) ont étendu les travaux de Bhattacherjee et Premkumar (2004) en ajoutant les variables de l'UTAUT (Venkatesh et al., 2003) à l'EDT. Ils montrent une réduction après usage sur les variables évaluées. Cependant, le nombre de recherches sur l'évolution de l'acceptabilité à l'acceptation reste limité.

Les rares travaux (Bhattacherjee & Premkumar, 2004 ; Karahanna et al., 1999 ; Venkatesh et al., 2011) sur l'évolution du jugement entre l'avant et l'après usage sont, à notre connaissance, basés uniquement sur le TAM (F. D. Davis, 1989) ou l'UTAUT (Venkatesh et al., 2003). Cependant, il est aujourd'hui reconnu que les utilisateurs perçoivent les produits au-delà de leurs qualités fonctionnelles (Bargas-Avila & Hornbaek, 2011). Ainsi, inclure d'autres variables que l'utilisabilité perçue comme le proposent Venkatesh et collaborateurs (2011), semble une approche intéressante pour comprendre les processus d'acceptabilité et d'acceptation. Dans la présente étude, l'EDT sera utilisée pour étudier l'évolution de l'acceptabilité à l'acceptation. L'originalité de ce travail se situe dans l'ajout de facteurs d'UX et de variables affectivo-motivationnelles dans l'évaluation avant et après usage.

2. Étude

L'adoption des produits technologiques est une problématique de recherche importante. Cette dernière se retrouve notamment pour les applications mobiles largement popularisées avec l'apparition des smartphones. Selon l'EDT, l'adoption d'un système commence avant le premier usage : s'intéresser au jugement avant usage paraît donc important dans le contexte des applications mobiles. En effet, avant de choisir d'installer une application depuis un magasin d'applications, les individus

sont dans l'impossibilité de tester l'application pour la juger. Or, à notre connaissance, le jugement avant usage dans ce contexte n'a pas été étudié ainsi que l'influence de l'UX et les variables affectivo-motivationnelles. Ainsi, dans cette étude, la différence entre le jugement avant et après usage a été évaluée sur les perceptions d'utilité, d'utilisabilité, de confiance, de stimulation, d'image de soi, de motivation intrinsèque et d'influence sociale.

De plus, afin de comprendre l'adoption des applications mobiles, le cadre théorique de l'EDT a été utilisé. En effet, l'adoption d'une application mobile est semblable aux situations décrites dans les travaux antécédents : lorsque les individus doivent sélectionner une application, ils n'ont accès qu'à peu d'informations comme la description de l'application ou des informations de seconde main comme les recommandations de l'entourage. À partir de ces informations, les personnes doivent former leurs jugements et leurs attitudes envers l'application. À partir du moment où ils installent et testent réellement l'application, ils sont en mesure de comparer leurs jugements a priori (i.e., jugement avant usage) avec leurs nouvelles représentations de l'application (i.e., jugement après usage). Selon les précédentes études, les utilisateurs décideraient alors de continuer ou non à utiliser l'application à partir de leurs attentes initiales et des confirmations/infirmations (i.e., les différences entre leurs attentes et la réalité).

2.1. Méthode

2.1.1. Matériel

L'application mobile grand public et multiplateformes de gestion du sommeil Sleep Better[42] a été sélectionnée (voir Figure 32). Ce matériel expérimental a été choisi pour plusieurs raisons. Tout d'abord, il devait pouvoir évoquer à la fois des qualités fonctionnelles et non fonctionnelles chez les participants. De plus, pour mobiliser un nombre important de participants dans des conditions naturelles, il était nécessaire que le système soit à la fois potentiellement utile, facile d'accès, conçu pour un usage répété mais encore assez peu diffusé. En effet, les participants devaient ne pas déjà utiliser l'application afin d'être en mesure d'évaluer le jugement avant

[42] Application développée par Runtastic disponible sous Android et iOS : https://www.runtastic.com/fr/apps/sleepbetter.

usage mais également que certains souhaitent tester l'application pour évaluer le jugement après usage chez les mêmes participants.

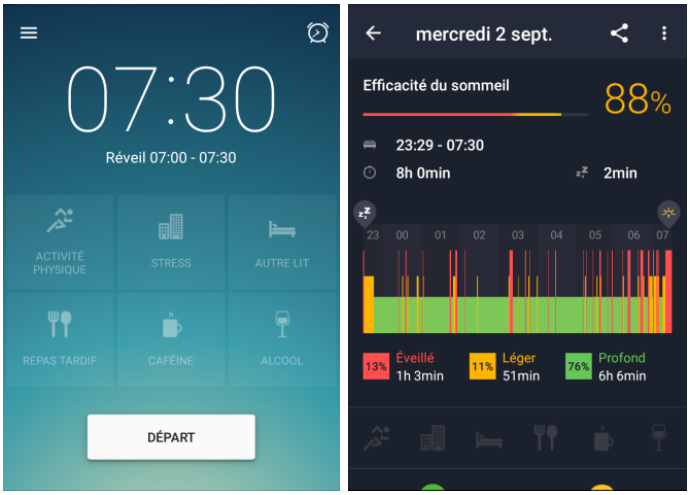


Figure 32. Captures d'écran de l'application Sleep Better

2.1.2. Mesures

Les mesures ne seront que brièvement présentées ci-dessous : pour une présentation détaillée de la construction des échelles, se reporter aux Chapitre 1 : Modélisation de l'adoption avant usage et Chapitre 2 : Jugement après usage. Ainsi, pour évaluer l'influence des facteurs d'UX et affectivo-motivationnels sur l'acceptabilité et l'acceptation, les dimensions suivantes ont été évaluées : l'Utilisabilité (4 items), l'Utilité (4 items), la Stimulation (4 items) et la Confiance (3 items) pour les facteurs UX ; la Motivation Intrinsèque (5 items), l'Image de Soi (4 items) et l'Influence Sociale (3 items) pour les facteurs affectivo-motivationnels et l'Intention d'Usage (4 items) pour la mesure de l'acceptabilité/acceptation.

Certains items ont été construits spécifiquement pour cette étude et d'autres sont basés sur des questionnaires existants comme l'UTAUT (Venkatesh et al., 2003), l'Attrakdiff (Hassenzahl et al., 2010) ou encore l'échelle d'engagement (*Engagement Scale)* (O'Brien & Toms, 2013). Pour chaque item, les participants devaient répondre sur une échelle de Likert en 11 points (entre 0 et 10). La position de chaque item dans le questionnaire a été rendue aléatoire afin de réduire les effets liés à l'ordre des questions (Benton & Daly, 1991). Enfin, selon l'EDT, l'intention de continuer à utiliser un produit est basée sur les représentations initiales (i.e., évaluations avant usage) et les infirmations (i.e., différences entre les représentations initiales et la réalité). Les

attentes sont mesurées en évaluant avant usage chaque dimension présentée ci-dessus. Les infirmations sont évaluées en calculant les différences entre l'évaluation avant et après usage pour chacune des dimensions.

2.1.3. Procédure

L'étude se déroulait en trois grandes étapes. Premièrement, les participants remplissaient le questionnaire avant usage après avoir lu une brève présentation de l'application (description et image – voir Annexe 7). Cette étape est identique à celle présentée dans le Chapitre 1 : Modélisation de l'adoption avant usage. Deuxièmement, à la fin du questionnaire, il leur était proposé de tester l'application et de participer à la suite de l'étude. Les participants qui choisissaient de tester l'application recevaient un mail leur expliquant comment installer l'application et leur donnant le lien vers la troisième étape correspondant au questionnaire après usage. Les participants avaient comme consigne de répondre au questionnaire après usage à la suite d'au minimum une utilisation de l'application. Les données entre ces deux questionnaires sont reliées de manière anonyme à l'aide d'un code unique (jour de naissance, les 3 premières lettres du prénom, numéro de rue et les 3 premières lettres de la ville de naissance).

2.1.4. Participants

318 participants ont rempli le questionnaire avant usage. Cependant, seulement 47 participants (39 femmes), n'ayant jamais testé l'application Sleep Better ont rempli le questionnaire avant et après usage. Ainsi, les analyses ont été conduites uniquement sur ces participants. Ces derniers ont été recrutés sur les réseaux sociaux : Facebook, Twitter et LinkedIn. L'âge moyen était de 29.19 ans (Écart-Type = 10.72).

3. Résultats

Afin de tester les effets de l'usage sur le jugement, des comparaisons entre l'évaluation avant et après usage ont été réalisées pour chaque dimension. Des modèles mixtes (Bates, Mächler, Bolker, & Walker, 2015 ; Gueorguieva & Krystal, 2004) ont été utilisés pour prendre en compte la non indépendance des données (Field, Miles, & Field, 2012) puisque les évaluations ont été complétées par les mêmes

participants[43]. Les analyses (voir Tableau 19) sur les données de cette étude montrent une diminution significative entre l'évaluation avant et après usage pour l'Intention d'Usage, l'Utilité, l'Image de Soi et l'Influence Sociale. En d'autres termes, les participants ont tendance à avoir un jugement significativement plus négatif sur ces dimensions.

Tableau 19. Statistiques descriptives (moyenne et écart-type) et comparaison statistique entre l'évaluation avant et après usage

Dimension	Avant Usage	Après usage	Différence
Intention Comportementale	6.59 (2.60)	4.89 (2.95)	χ^2 (1) = 17.00, p < .001
Utilité	6.74 (2.19)	5.13 (2.41)	χ^2 (1) = 25.57, p < .001
Utilisabilité	7.58 (1.61)	7.73 (1.95)	χ^2 (1) =.36, NS
Stimulation	6.80 (2.24)	6.55 (1.74)	χ^2 (1) = .72, NS
Confiance	5.56 (1.96)	4.96 (2.40)	χ^2 (1) = 3.42, NS
Image de Soi	3.77 (2.03)	2.73 (1.96)	χ^2 (1) = 14.58, p < .001
Motivation Intrinsèque	5.54 (2.19)	5.18 (1.87)	χ^2 (1) = 1.47, NS
Influence Sociale	4.94 (2.32)	2.75 (2.29)	χ^2 (1) = 33.29, p < .001

[43] Les modèles mixtes (Baayen, Davidson, & Bates, 2008 ; Bates, Mächler, Bolker, & Walker, 2015) ont été sélectionnés parce qu'ils offrent davantage de flexibilité et de robustesse dans la modélisation que les ANOVA à mesures répétées (Gueorguieva & Krystal, 2004). Afin de tester les effets d'une variable, des modèles emboités (*nested models*) sont comparés. Techniquement, deux modèles sont comparés pour chaque effet testé (voir pour exemple Baayen et al., 2008) : un premier modèle (i.e., modèle nul) incluant uniquement l'effet aléatoire (pour prendre en compte le lien entre les données récoltées auprès d'un même participant) et un second incluant la variable testée et l'effet aléatoire (modèle plus complexe – modèle alternatif). Pour évaluer la contribution de chaque variable, la différence de déviance (test du χ^2) est évaluée entre les deux modèles emboités avec le seuil de significativité α = .05. Dans notre cas, si le modèle alternatif incluant la variable temps réduit significativement la déviance (p < .05) il est considéré comme meilleur (i.e., la variable indépendante temps a une influence significative sur la variable dépendante) et on considère qu'il y a une différence entre l'avant et l'après usage.

Selon l'EDT (Bhattacherjee & Premkumar, 2004 ; Venkatesh et al., 2011), la poursuite de l'usage d'un produit est basée sur les représentations initiales (i.e., évaluation avant usage) et les infirmations (i.e., différence entre les représentations initiales et la réalité). Afin de tester ces hypothèses, une régression multiple sur l'intention comportementale après usage a été estimée. Les résultats (voir Figure 33) indiquent une influence significative des représentations initiales d'Utilité (β = .85, p < .01), d'Utilisabilité (β = .55, p < .01) et de Stimulation (β = -.49, p < .05). On retrouve aussi une influence des infirmations pour l'Utilité (β = .50, p < .05), l'Utilisabilité (β = .36, p < .05), la Stimulation (β = -.47, p < .05) et la Confiance (β = .35, p < .05). Enfin, en ce qui concerne les facteurs affectivo-motivationnels, seule la Motivation Intrinsèque (β = .68, p < .01) semble avoir un effet sur l'intention comportementale après usage.

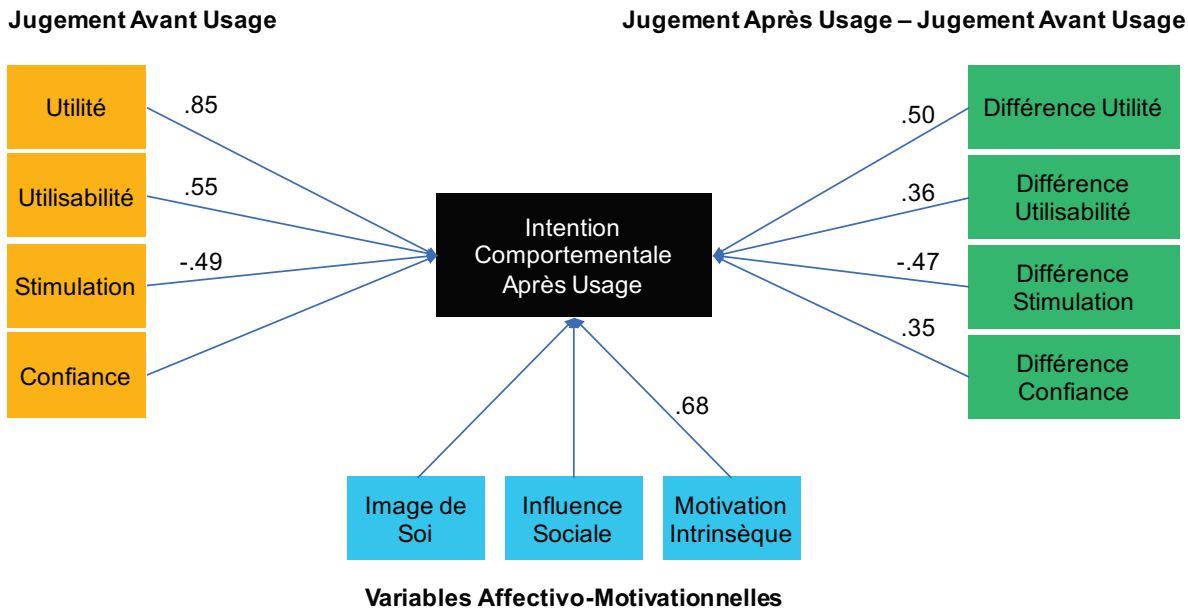


Figure 33. Résultats de la régression multiple mettant en œuvre notre proposition de modèle EDT[44, 45]

[44] Seuls les coefficients bêta significatifs (p < .05) sont présentés.

[45] Les variables intitulées « Différence […] » correspondent aux variables calculées en réalisant la différence entre le jugement après usage et le jugement avant usage. Par conséquent, un coefficient positif indique que plus la différence est importante (plus l'écart

En résumé, selon les résultats obtenus, plus l'application est perçue comme utile et utilisable avant usage, plus les participants auraient tendance à vouloir poursuivre leur utilisation. De plus, plus la différence de perception entre l'avant et l'après usage en termes d'Utilité, d'Utilisabilité et de Confiance est importante, moins les participants souhaitent poursuivre leur usage. De plus, plus les participants sont motivés intrinsèquement à utiliser l'application, plus ils souhaitent poursuivre leur usage. A l'inverse, la stimulation perçue aurait un effet négatif. De plus, deux modèles expliquant l'intention comportementale ont été comparés afin de tester l'intérêt d'ajouter des variables UX et affectivo-motivationnelles au modèle EDT en plus de celles du TAM. Le premier modèle inclut uniquement les variables du TAM comme dans l'étude de Bhattacherjee et Premkumar (2004). Le second modèle inclut les variables UX et affectivo-motivationnelles en plus de celles du TAM. Les analyses indiquent que notre modèle explique significativement ($F(7, 47) = 2.69$, $p < .05$) mieux l'intention comportementale qu'un modèle incluant uniquement les variables du TAM : $R^2 = .67$ *vs*. .74.

4. Discussion et conclusion

Les résultats montrent que l'usage pourrait avoir une influence négative et significative sur l'Intention Comportementale, l'Utilité, l'Utilisabilité, la Confiance, l'Image de Soi et l'Influence Sociale dans le contexte d'une application mobile. Cela est en accord avec de précédentes études (Bhattacherjee & Premkumar, 2004 ; Venkatesh et al., 2011). De plus, afin de cerner au mieux le jugement envers l'application mobile Sleep Better, la perception des qualités fonctionnelles et non fonctionnelles, mais également des facteurs affectivo-motivationnels a été intégrée au cadre théorique de l'EDT. Ainsi, l'influence du jugement a priori (i.e., jugement avant usage) et des infirmations sur l'intention de continuer à utiliser le produit a été évaluée. Trois principaux résultats sont à retenir : premièrement, les attentes en termes d'Utilité, d'Utilisabilité et de Stimulation influencent significativement l'Intention de continuer à utiliser. Deuxièmement, les infirmations (i.e., différence entre les perceptions initiales et la réalité) d'Utilité, d'Utilisabilité, de Stimulation et de Confiance sont des prédicteurs

entre l'avant et l'usage est important), plus l'influence de la variable considérée est importante.

significatifs de l'Intention de continuer à utiliser. Enfin, la Motivation Intrinsèque semble être le seul facteur affectivo-motivationnel prédicteur de l'Intention Comportementale. En d'autres termes, les perceptions initiales de qualités ainsi que les différences entre ces perceptions et la réalité influencent l'intention de poursuivre l'usage d'un produit. Ainsi, il semble intéressant de prendre en compte le jugement avant usage en évaluant la perception a priori des qualités fonctionnelles et non fonctionnelles ainsi que des facteurs affectivo-motivationnels. Cela pourrait permettre de mieux comprendre l'adoption des produits et la poursuite de l'usage au cours du temps.

Néanmoins, plusieurs limites peuvent être soulevées. Tout d'abord, l'échantillon est relativement réduit. Afin de confirmer les résultats, il est nécessaire de conduire de nouvelles études sur des échantillons plus importants. Cela permettrait notamment d'utiliser des SEMs pour étudier les liens entres les variables et leur évolution au cours du temps. Deuxièmement, une seule application possédant de bonnes qualités fonctionnelles, a été proposée aux participants. En effet, toutes les dimensions rattachées aux qualités fonctionnelles (i.e., Utilité, Utilisabilité, Stimulation et Confiance) sont jugées relativement positivement avec des moyennes au-dessus de 5.5 (sur 10). Afin de mieux comprendre l'évolution du jugement, il serait souhaitable de mener d'autres études avec des produits possédant des qualités perçues plus variées. Enfin, les participants n'ont évalué l'application qu'à deux reprises : une fois avant usage et une fois après usage. Afin d'avoir un point de vue plus complet sur l'évolution de l'acceptabilité à l'acceptation, il apparait primordial de réaliser des études longitudinales à plus long terme avec davantage de points de mesures.

Chapitre 4 : Construction des échelles à items uniques pour la mesure avant et après usage[46]

1. Introduction

Lors de la création des échelles d'évaluation du jugement utilisateur et de ses composantes, la longueur des questionnaires est rapidement apparue comme un obstacle. En effet, selon la littérature, de nombreuses dimensions pourraient avoir une influence sur le jugement envers un système. Par conséquent, un nombre important de variables est nécessaire pour les mesurer et ainsi obtenir une représentation la plus complète possible. Au cours des recueils de données de cette thèse, la longueur des questionnaires utilisés a été critiquée à de nombreuses reprises par les participants, bien qu'elle soit similaire à la littérature (i.e., la longueur maximale du questionnaire était de moins de 60 items). Il semblait, de ce fait, important d'explorer des solutions pour faciliter les passations sans pour autant limiter les qualités psychométriques et les possibilités de modélisations statistiques. Cet objectif était d'autant plus important que cette thèse s'intégrait dans un projet avec un objectif appliqué. Ainsi, permettre et faciliter l'exploitation des travaux réalisés faisait partie intégrante des missions à conduire au cours de la thèse. Ainsi, dans ce chapitre, il sera tout d'abord présenté les échelles à items uniques sur le plan théorique et les principes statistiques sous-jacents. Enfin, la création et l'étude psychométrique des échelles à items uniques seront abordées.

2. Les échelles à items uniques

En psychologie, les travaux font généralement appel à des concepts théoriques difficiles à mesurer directement (Groth-Marnat, 2003). Pour cela, des items peuvent notamment être utilisés afin d'évaluer des manifestations observables du construit

[46] Une partie de cette étude a été publiée dans la conférence Interaction-Homme Machine (IHM'16) à Genève (Martin, Jamet, Erhel, & Rouxel, 2016a).

théorique (DeVellis, 2017) et d'estimer les variables latentes associées (Bollen, 2002) (voir ci-après pour une explication plus détaillée). Par exemple, l'évaluation directe de l'anxiété est pratiquement impossible. Par conséquent, ce sont généralement des manifestations observables de la dimension évaluée qui sont utilisées. Dans le cadre de l'anxiété, cela pourrait être un comportement (e.g., « en période d'examen, je me ronge plus souvent les ongles ») qui est mesuré. En d'autres termes, le concept théorique n'étant pas mesurable directement, il est nécessaire d'obtenir le score le plus proche possible du score vrai (i.e., ce que l'on souhaite réellement mesurer). Ensuite, des variables latentes sont généralement construites, par l'intermédiaire de modèles statistiques (e.g., SEM), pour représenter les dimensions théoriques à partir des données récoltées. Ainsi, lorsque plusieurs dimensions psychologiques sont évaluées, les participants peuvent avoir à compléter de longs questionnaires, potentiellement fastidieux et ennuyeux (Robins, Hendin, & Trzesniewski, 2001). De ce fait, l'usage d'échelles courtes et notamment d'échelles à items uniques (i.e., un item par dimension théorique) s'est développé (e.g., Abdel-Khalek, 2006 ; Robins et al., 2001). Ces questionnaires permettent de réduire les temps de réponses et par conséquent de faciliter la conduite d'expérimentations (Christophersen & Konradt, 2011). Cependant, l'exploitation des données extraites de ces questionnaires comporte plusieurs verrous. Ces limites sont principalement liées aux principes sous-jacents à la modélisation avec des variables latentes qui seront présentées au sein de la partie suivante.

3. Principes statistiques sous-jacents aux échelles à items uniques

Avant de préciser la procédure et les difficultés inhérentes à la modélisation avec des items uniques, il parait important de présenter le principe des variables latentes basées sur des échelles à items multiples (i.e., condition classique de l'utilisation des SEMs). Comme expliqué précédemment, les variables latentes sont utilisées pour représenter des concepts théoriques non directement mesurables comme une variable psychologique. En d'autres termes (voir Figure 34), le score vrai (t) n'est pas directement accessible, et ce sont généralement les manifestations observables (DeVellis, 2017) de ces dimensions qui sont utilisées comme

approximation. Les manifestations observables (y) sont considérées comme le résultat du score vrai et de l'erreur de mesure (e) (Alagumalai & Curtis, 2005).

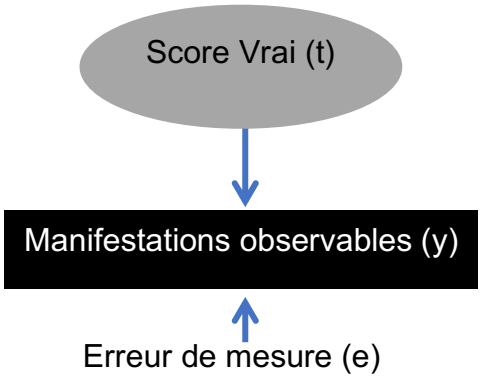


Figure 34. Score vrai, manifestations observables et erreur de mesure

Ainsi, une variable latente est construite statistiquement comme la variable à l'origine des covariances observées entre les items, outre les erreurs de mesure (δ) (Beaujean, 2014) comme présenté dans la Figure 35.

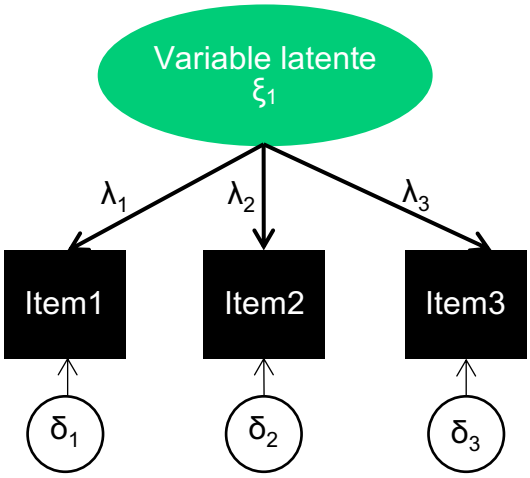


Figure 35. Modèle de mesure avec une variable latente et 3 items[47]

Cependant, ces variables latentes requièrent plusieurs variables observables pour être estimées. En effet, un modèle structural doit être identifié pour être estimable (voir Kenny & Milan (2012) pour une revue sur l'identification des SEMs). Le nombre

[47] Pour des raisons de standardisation (Beaujean, 2014), soit la variance de la variable latente, soit la saturation sur un item doit être fixée à 1. Ici, le lien (λ_1) entre la variable latente et l'item1 a été fixé à 1.

d'informations non redondantes (i.e., variables observables) doit être égal ou supérieur au nombre de paramètres à estimer (Beaujean, 2014). Ainsi, par exemple, dans le cadre d'une modélisation avec une seule variable latente, au moins trois indicateurs sont nécessaires (voir Figure 35). En effet, 6 informations[48] sont disponibles :

- Variance de l'item 1 : VAR(Item1)
- Variance de l'item 2 : VAR(Item2)
- Variance de l'item 3 : VAR(Item3)
- Covariance entre l'item 1 et l'item 2 : COV(Item1, Item2)
- Covariance entre l'item 2 et l'item 3 : COV(Item2, Item3)
- Covariance entre l'item 1 et l'item 3 : COV(Item1, Item3)

Et 6 paramètres sont à calculer[47] :

- Variance de la variable latente : VAR(ξ_1)
- Erreur de mesure de l'item 1 : δ_1
- Erreur de mesure de l'item 2 : δ_2
- Erreur de mesure de l'item 3 : δ_3
- Saturation de l'item 2 : λ_2
- Saturation de l'item 3 : λ_3

Par conséquent, le modèle de mesure est juste identifié (i.e., le nombre d'informations non redondantes est égal au nombre de paramètres) et peut être estimé. À l'inverse, lorsqu'une variable latente est mesurée uniquement par un item (voir Figure 36), le modèle devient non identifié (i.e., le nombre d'informations disponibles est inférieur au nombre de paramètres) et par conséquent non estimable directement.

[48] Il est possible de calculer le nombre de variances et covariances (i.e., informations non redondantes) à l'aide de la formule suivante : $\frac{p(p+1)}{2}$ où p correspond au nombre de variables observables.

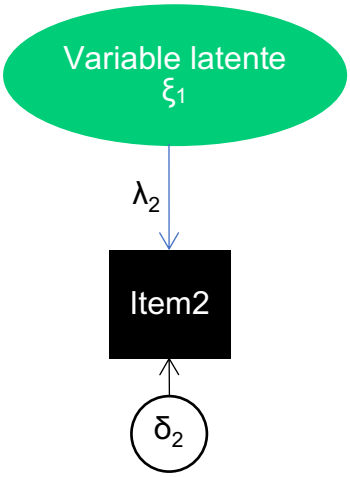


Figure 36. Modèle de mesure avec une variable latente et un item : modèle non identifié et non estimable

En effet, une seule information non redondante est disponible : la variance de l'item 2 (VAR(Item2)) et trois paramètres sont à calculer : la variance de la variable latente (VAR(ξ_1)), la saturation entre la variable latente et l'item 2 (λ_2) ainsi que le terme d'erreur de mesure de l'item 2 (δ_2). Ainsi, pour être en mesure d'estimer les paramètres d'un modèle avec une variable latente et un seul item, plusieurs modifications dans la modélisation sont nécessaires.

L'approche la plus simple consiste à réaliser les modifications comme présentées en Figure 37. Premièrement, la saturation (*factor loading* - λ) entre la variable latente et l'item est fixée à 1 (Kline, 2010). Deuxièmement, l'erreur de mesure (*measurement error variance* - δ) de l'item est fixée à 0. Cette configuration revient à considérer que l'item mesure parfaitement la variable latente (Hayduk & Littvay, 2012). Or, cette hypothèse de mesure pure est difficile à soutenir et peut conduire à des erreurs dans la modélisation (Cole & Preacher, 2014). Pour faire face à ce problème, une autre approche consiste à fixer l'erreur de mesure à une valeur spécifique non nulle (Beaujean, 2014 ; Kline, 2010). Cette valeur peut être déterminée de plusieurs manières. Lorsque des items uniques sont utilisés depuis la création de l'échelle, l'erreur de mesure est généralement fixée à partir de la fiabilité attendue (*expected reliability*) et de la variance de l'item (Fuchs & Diamantopoulos, 2009)[49]. Lorsque la

[49] Afin de calculer la variance de l'erreur de mesure, cette formule est généralement utilisée (Fuchs & Diamantopoulos, 2009) : $\varepsilon = (1 - Expected\ Reliability) * Var(Item)$.

réduction du nombre d'items est effectuée après une analyse factorielle sur une échelle à items multiples, les erreurs de mesure peuvent être fixées à partir des erreurs de mesure estimées au cours de cette analyse factorielle (Hayduk & Littvay, 2012 ; Oberski & Satorra, 2013). Dans le cadre de ces travaux de thèse, la deuxième option a été testée. En d'autres termes, l'approche sélectionnée consiste à récupérer les valeurs d'erreurs de mesure obtenues lors des analyses factorielles pour les réutiliser afin de fixer les erreurs de mesure à une valeur non nulle dans les échelles à items uniques. Pour s'assurer que les erreurs de mesure estimées au sein d'un échantillon restent stables au travers des expérimentations, des données issues de plusieurs recueils seront utilisées.

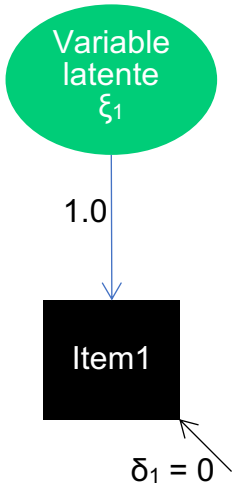


Figure 37. Variable latente avec un seul item sans prise en compte de l'erreur de mesure : modèle identifié et estimable

4. Création des échelles de mesure avant et après usage à items uniques

Une version avant usage et une version après usage des échelles de mesures ont été construites. Afin de tester les méthodes proposées, des expérimentations ont été conduites à partir des échantillons utilisés au sein des chapitres précédents. En effet, comme exposé ci-dessus, la méthode sélectionnée pour la prise en compte des erreurs de mesure repose sur l'utilisation de valeurs non nulles pour chaque erreur de mesure. Pour ce faire, les données récoltées ainsi que les résultats d'analyses factorielles présentés dans les chapitres précédents ont été réutilisés dans ce chapitre.

Afin de tester l'efficacité de cette approche, les résultats obtenus avec des questionnaires à items multiples ont été comparés à ceux obtenus en gardant uniquement un item par dimension (équivalent à des échelles à items uniques) en incluant (méthode « correction ») ou non (méthode « naïve ») les erreurs de mesure.

4.1. Échelles avant usage à items uniques[50]

La construction des échelles à items uniques repose sur les résultats des analyses factorielles réalisées lors des études psychométriques (voir Chapitre 1 : Modélisation de l'adoption avant usage). Cependant, afin de s'assurer de la stabilité des paramètres des modèles (Fan, 2003), les données ont été de nouveau analysées en rajoutant une méthode *bootstrap* (1000 échantillons). Cette méthode permet de ré-échantillonner (tirages aléatoires avec remise) les données (Bone, Sharma, & Shimp, 1989). De plus, les données issues des items ont été standardisées avant d'être utilisées dans le modèle. Pour rappel, les données se composent de deux échantillons indépendants : un premier groupe de 305 participants ayant évalué avant usage l'application Sleep Better et un second groupe de 247 participants ayant évalué avant usage l'application Google Fit.

Sur la base des résultats obtenus à partir des analyses factorielles, un item a été sélectionné par dimension évaluée. Cette sélection repose sur des indices statistiques (i.e., saturations les plus élevées et erreurs de mesure les plus faibles) ainsi que sur la formulation des questions (i.e., les questions les plus simples ont été choisies). Les items suivants ont été choisis pour les échelles avant usage :

- ***Utilité*** : Je trouve que cette application pourrait m'être utile.
- ***Utilisabilité*** : L'utilisation de cette application semble simple.
- ***Stimulation*** : Cette application me semble novatrice.
- ***Confiance*** : J'ai confiance dans les informations que cette application pourrait fournir.

[50] Les données utilisées sont identiques à celle présentées dans les analyses psychométriques des échelles avant usage sur l'échantillon ayant évalué l'application Sleep Better et Google Fit. Une partie de cette étude a été publiée dans la conférence Interaction-Homme Machine (IHM'16) à Genève (Martin et al., 2016a).

- ***Image de Soi*** : Je pourrais être perçu(e) positivement par les autres si j'utilisais cette application.
- ***Motivation Intrinsèque*** : J'utiliserai cette application parce que je pourrais avoir beaucoup de plaisir à interagir avec cette application.
- ***Influence Sociale*** : Les personnes de mon entourage qui comptent pour moi pourraient m'encourager à utiliser cette application.

Pour chacun de ces items, les erreurs de mesure estimées au cours des analyses factorielles par *bootstrapping* ont été extraites. Ces paramètres sont les suivants (voir Annexe 17 pour l'ensemble des paramètres) :

- ***Utilité*** = .10
- ***Utilisabilité*** = .18
- ***Stimulation*** = .18
- ***Confiance*** = .18
- ***Image de Soi*** = .21
- ***Motivation Intrinsèque*** = .31
- ***Influence Sociale*** = .24

Afin de s'assurer de l'intérêt de cette approche, il paraissait nécessaire de comparer les résultats obtenus à l'aide des échelles à items uniques (incluant ou non les erreurs de mesure) et multiples. Afin d'examiner les éventuelles différences, les données collectées avec les échelles à items multiples ont servi de référence. Elles ont été comparées aux données obtenues avec les échelles à items uniques avec la prise en compte de l'erreur de mesure puis aux données collectées avec les échelles à items uniques sans la méthode proposée ici. Pour cela, des scores factoriels ont été calculés à partir des données collectées avec ces 3 échelles. Puis des corrélations ont été calculées entre les données issues des échelles à items multiples (16 items [4 dimensions] pour l'échelle UX et 12 items [3 dimensions] pour l'échelle A-M) et les données issues des échelles à items uniques (4 items [4 dimensions] pour l'échelle UX et 3 items [3 dimensions] pour l'échelle A-M) incluant ou non la prise en compte des erreurs de mesure.

Les résultats sont présentés dans le Tableau 20 pour l'échelle UX et dans le Tableau 21 pour l'échelle A-M. Plus précisément, il a d'abord été calculé la corrélation

entre les scores factoriels obtenus avec l'échelle à items multiples et les scores factoriels obtenus avec l'échelle à items uniques prenant ou non en compte l'erreur de mesure estimée dans le premier échantillon (305 participants ayant évalué l'application Sleep Better). Puis, sur les données collectées dans le second échantillon (247 participants qui ont évalué l'application Google Fit) une corrélation a été estimée entre les scores factoriels obtenus avec l'échelle à items multiples et les scores factoriels obtenus avec l'échelle à items uniques prenant ou non en compte l'erreur de mesure estimée dans le premier échantillon (305 participants ayant évalués application Sleep Better). Cette dernière approche permet de s'assurer que des erreurs estimées dans un premier échantillon pourront potentiellement apporter une amélioration lors des utilisations futures des échelles à items uniques.

Tableau 20. Comparaisons entre les échelles avant usage UX à items multiples, à items uniques corrigés et à items uniques[51]

Utilité				**Utilisabilité**				**Stimulation**			
C1	NC1	C2	NC2	C1	NC1	C2	NC2	C1	NC1	C2	NC2
.971	.966	.953	.945	.949	.945	.906	.902	.941	.933	.935	.924

Confiance			
C1	NC1	C2	NC2
.949	.929	.955	.949

Tableau 21. Comparaisons entre les échelles avant usage A-M à items multiples, à items uniques corrigés et à items uniques[51]

Influence Sociale				**Image de Soi**				**Motivation**			
C1	NC1	C2	NC2	C1	NC1	C2	NC2	C1	NC1	C2	NC2
.928	.914	.954	.943	.920	.878	.952	.939	.912	.899	.920	.887

Sur le plan descriptif, les corrélations entre les scores factoriels obtenus avec les échelles à items multiples et ceux obtenus avec les échelles à items uniques ont tendance être plus faibles lorsque les erreurs de mesure ne sont pas prises en compte.

[51] C1 : Corrélation sur l'échantillon 1 (évaluation de l'application Sleep Better) entre les scores factoriels estimés à partir de l'échelle à items multiples et les scores factoriels estimés à partir de l'échelle à items uniques en incluant les erreurs de mesure.

NC1 : Corrélation sur l'échantillon 1 (évaluation de l'application Sleep Better) entre les scores factoriels estimés à partir de l'échelle à items multiples et les scores factoriels estimés à partir de l'échelle à items uniques sans les erreurs de mesure.

C2 : Corrélation sur l'échantillon 2 (évaluation de l'application Google Fit) entre les scores factoriels estimés à partir de l'échelle à items multiples et les scores factoriels estimés à partir de l'échelle à items uniques en incluant les erreurs de mesure estimées dans l'échantillon 1.

NC2 : Corrélation sur l'échantillon 2 (évaluation de l'application Google Fit) entre les scores factoriels estimés à partir de l'échelle à items multiples et les scores factoriels estimés à partir de l'échelle à items uniques sans les erreurs de mesure.

De plus, la prise en compte des erreurs de mesure, même lorsque ces dernières ont été estimées dans un autre échantillon, semble améliorer les scores factoriels.

4.2. Échelles après usage à items uniques[52]

En ce qui concerne les échelles après usage à items uniques, au regard des données collectées, les analyses n'ont pu être conduites qu'au sein d'un seul échantillon. Ainsi, la construction des échelles après usage à items uniques est basée sur les résultats des analyses factorielles réalisées lors des études psychométriques : voir Chapitre 2 : Jugement après usage. Comme pour les échelles avant usage, la méthode de ré-échantillonnage *bootstrap* (1000 échantillons) a été utilisée afin de s'assurer de la stabilité des paramètres des modèles (Fan, 2003).

Afin de construire l'échelle, un item par dimension évaluée a été choisi en utilisant des indices statistiques (saturations les plus élevées et erreurs de mesure les plus faibles) ainsi que sur la formulation des questions. Les items suivants ont été choisis :

- ***Utilité*** : Je trouve que cette application m'est utile.
- ***Utilisabilité*** : Se servir de cette application est facile.
- ***Stimulation*** : Les fonctions proposées par cette application sont innovantes.
- ***Confiance*** : Les informations apportées par cette application semblent de bonne qualité.
- ***Esthétisme*** : Cette application est esthétiquement attrayante.
- ***Absorption*** : Je me sens, à certains moments, complètement absorbé(e) par l'interaction avec cette application.
- ***Émotions Positives*** : Quand j'utilise cette application, je me sens content(e), satisfait(e).
- ***Émotions Négatives*** : Quand j'utilise cette application, je me sens agacé(e), énervé(e), exaspéré(e).
- ***Habitudes*** : Utiliser cette application est devenu automatique pour moi.

[52] Les données utilisées sont identiques à celles présentées dans les analyses psychométriques des échelles après usage sur l'échantillon ayant évalué l'application Waze.

- ***Image de Soi*** : L'utilisation de cette application renvoie une image positive de moi.
- ***Motivation Intrinsèque*** : J'utilise cette application parce qu'elle m'amuse beaucoup.
- ***Influence Sociale*** : Les personnes de mon entourage que j'ai l'habitude d'écouter me conseillent d'utiliser cette application.

Pour chacun de ces items, les erreurs de mesure estimées au cours des analyses factorielles par *bootstrapping* ont été extraites. Ces paramètres sont les suivants :

- ***Utilité*** = .26
- ***Utilisabilité*** = .15
- ***Stimulation*** = .26
- ***Confiance*** = .13
- ***Esthétisme*** = .24
- ***Absorption*** = .42
- ***Émotions Positives*** = .43
- ***Émotions Négatives*** = .23
- ***Habitudes*** = .18
- ***Image de Soi*** = .30
- ***Motivation Intrinsèque*** = .42
- ***Influence Sociale*** = .20

Identiquement à ce qui a été réalisé pour le questionnaire avant usage, les scores factoriels obtenus avec les échelles à items multiples (18 items [5 dimensions] pour l'échelle UX et 29 items [7 dimensions] pour l'échelle A-M), à items uniques (5 items [5 dimensions] pour l'échelle UX et 7 items [7 dimensions] pour l'échelle A-M) incluant ou non les erreurs de mesure ont été comparés. Pour ce faire, les corrélations entre les scores obtenus avec les échelles à items multiples et les échelles à items uniques, en incluant les erreurs de mesure ou non, ont été calculées : voir le Tableau 22 pour l'échelle UX et le

Tableau 23 pour l'échelle A-M.

Tableau 22. Corrélations entre les échelles UX après usage à items multiples, à items uniques corrigés et à items uniques[53]

Utilité		Utilisabilité		Stimulation		Confiance		Esthétisme	
C	NC	C	NC	C	NC	C	NC	C	NC
.974	.910	.963	.950	.911	.910	.963	.966	.931	.931

Tableau 23. Corrélations entre les échelles A-M après usage à items multiples, à items uniques corrigés et à items uniques[53]

Influence Sociale		Image de Soi		Motivation		Habitudes	
C	NC	C	NC	C	NC	C	NC
.947	.943	.933	.896	.914	.829	.939	.935
Emotions Négatives		**Emotions Positives**		**Absorption**			
C	NC	C	NC	C	NC		
.915	.915	.916	.819	.900	.836		

Les corrélations semblent plus élevées lorsque les erreurs de mesure sont prises en compte lors de l'estimation des scores factoriels.

5. Discussion et Conclusion

Cette étude présente les premières étapes de développement d'échelles courtes basées sur des items uniques. Avant la création de ces échelles, une première version des questionnaires avec plusieurs items par dimension évaluée avait été créée. Bien que ces échelles présentent des qualités psychométriques satisfaisantes,

[53] C : Corrélation entre les scores factoriels estimés à partir de l'échelle à items multiples et les scores factoriels estimés à partir de l'échelle à items uniques en incluant les erreurs de mesure sur l'échantillon ayant évalué l'application Waze.

NC : Corrélation entre les scores factoriels estimés à partir de l'échelle à items multiples et les scores factoriels estimés à partir de l'échelle à items uniques sans les erreurs de mesure sur l'échantillon ayant évalué l'application Waze.

la longueur des échelles est rapidement apparue comme un obstacle majeur à la réalisation d'expérimentations. En effet, lors du recueil de données avec les échelles à items multiples, de très nombreux participants se sont plaints de la similarité entre les items ainsi que du nombre d'items à compléter. De plus, une partie de l'échantillon a abandonné durant la complétion du questionnaire, ce qui pourrait en partie être imputé à sa longueur.

Les problèmes liés à la longueur des questionnaires sont bien connus des chercheurs et ont fait l'objet de nombreuses publications. Ainsi, les participants percevraient la redondance des items (liée aux faibles variations entre les items évaluant une même dimension) comme un indicateur de faible validité de l'étude (e.g., Christophersen & Konradt, 2011). D'autres études ont montré que l'augmentation du nombre d'items pouvait générer des données manquantes (e.g., Dolbier, Webster, McCalister, Mallon, & Steinhardt, 2005). Ainsi, pour faciliter la conduite d'expériences, il est apparu pertinent de développer de nouvelles échelles courtes. Pour cela, un item par dimension évaluée a été sélectionné sur des critères statistiques et à partir de la formulation des questions. De plus, l'erreur de mesure associée à chaque item a été extraite des analyses factorielles. Ensuite, les scores obtenus avec les échelles à items multiples et uniques ont été comparés. Ainsi, selon les corrélations estimées entre les scores factoriels avec des échelles à items multiples et uniques avec ou sans prise en compte de des erreurs de mesure, la prise en compte de l'erreur de mesure semble améliorer les résultats (i.e., les scores factoriels semblent plus corrélés lorsque les erreurs de mesure sont prises en compte dans les échelles à items uniques). Ces différences se retrouvent à la fois dans les échantillons avant et après usage.

De plus, dans l'échantillon avant usage, il a été possible de tester l'approche dans des conditions plus proches d'un futur usage réaliste des questionnaires. En effet, des scores factoriels ont été estimés sur un échantillon en prenant en compte les erreurs de mesure estimées dans un autre échantillon. Cette approche correspondrait à l'usage d'un questionnaire à items uniques où la prise en compte des erreurs de mesure aurait lieu lors de l'analyse des données. Les résultats ont montré l'apport de la méthode.

Plusieurs limites et perspectives d'améliorations sont à noter. Premièrement, l'approche proposée pour estimer les erreurs de mesure présente un désavantage important. En effet, avant d'être en mesure d'utiliser des échelles courtes, il est nécessaire de récolter des données avec des échelles avec plusieurs items par

dimension évaluée afin de pouvoir estimer les erreurs de mesure. De plus, comme les erreurs de mesure ont été estimées uniquement dans un échantillon, les paramètres estimés peuvent être biaisés. Afin d'évaluer ce potentiel biais, des futurs travaux devront se concentrer sur la collecte de données auprès de plusieurs échantillons et sur plusieurs produits avec des échelles à items multiples. Ces données permettront de s'assurer de l'invariance des erreurs de mesure. De plus, ces données pourraient permettre d'affiner les résultats et d'estimer une erreur moyenne pour chaque item.

Deuxièmement, la méthode employée pour prendre en compte les erreurs de mesure est basée sur des valeurs estimées et moyennées au sein de plusieurs échantillons. Par conséquent, ces valeurs sont fixes (variances exactes du terme d'erreur de mesure) et correspondent à des estimations issues de la distribution exacte de la variable (i.e., valeur vraie). Ainsi, une incertitude existe autour de la valeur estimée (Oberski & Satorra, 2013). Certains auteurs (e.g., Oberski & Satorra, 2013) ont proposé l'introduction de cette incertitude dans les modélisations et montrent une amélioration des estimations. Cependant, ces méthodes restent difficiles à implémenter. Troisièmement, l'approche bayésienne pourrait potentiellement être utilisée pour estimer les erreurs de mesure. En effet, cette méthode permet d'intégrer lors de l'estimation des informations préalables (i.e., distribution des variables a priori ; Lynch, 2007). Appliquée au terme d'erreur, l'approche bayésienne devrait permettre de prendre en compte les données estimées préalablement (distributions a priori des erreurs estimées à l'aide des échelles à items multiples) mais également d'ajuster ces valeurs aux données du nouvel échantillon utilisé. Néanmoins, de la même manière que pour l'intégration de l'incertitude, l'approche bayésienne comporte aujourd'hui de nombreuses limites d'implémentation. En effet, les estimations requièrent beaucoup plus de temps et de puissance de calcul. De plus, les risques de non convergence des modèles peuvent être importants et difficiles à gérer. En résumé, ces deux méthodes semblent théoriquement intéressantes. Toutefois, elles comportent encore aujourd'hui de nombreuses limites empêchant une utilisation par les praticiens.

Enfin, pour conclure ce travail préliminaire, réduire le nombre d'items est un sujet d'intérêt pour les chercheurs et les praticiens. Cependant, la simple réduction du nombre d'items sans méthode corrective peut amener à des erreurs dans les modélisations (Cole & Preacher, 2014). Ainsi, au sein de cette étude, une approche a été explorée pour réduire le nombre d'items tout en limitant les problèmes liés aux erreurs de mesure et montrent des résultats intéressants. Cette approche pourrait

faciliter les expérimentations (Christophersen & Konradt, 2011) notamment longitudinales, réduire les données manquantes dues aux abandons au cours de l'étude et offrir de la flexibilité sur le plan de la modélisation statistique puisque le nombre de paramètres à estimer se trouve réduit (Oberski & Satorra, 2013).

Chapitre 5 : Preuve de concept - *Valeurs seuils et librairie R*

Les questionnaires sont des outils très populaires auprès des chercheurs et des praticiens. Cependant, l'interprétation des résultats issus de ceux-ci peut s'avérer difficile (Bangor, Kortum, & Miller, 2009, 2008). En effet, une partie de l'analyse des questionnaires repose sur des statistiques descriptives ou des graphiques (Lallemand & Gronier, 2015). Afin de faciliter l'exploitation des échelles construites dans cette thèse, une approche basée sur des régressions logistiques a été testée dans une perspective exploratoire. De plus, une librairie R a été développée pour implémenter les diverses pistes explorées dans cette thèse[54] et sera présentée dans ce chapitre.

1. Création des valeurs seuils pour les questionnaires

La détermination de valeurs seuils pour l'interprétation des questionnaires est une demande récurrente des expérimentateurs utilisant des questionnaires dans leurs études. En effet, les questionnaires renvoient des résultats qu'il convient ensuite de remplacer par rapport à une moyenne, un seuil permettant alors de qualifier le résultat obtenu. Les valeurs numériques récupérées par l'intermédiaire des échelles sont en effet rarement directement interprétables (Bangor et al., 2009, 2008). Dans le cadre de l'évaluation de produits technologiques, la comparaison de plusieurs systèmes est relativement aisée à l'aide d'outils statistiques simples (e.g., t de Student). Cependant, lorsqu'un seul système est évalué (e.g., lorsqu'aucun produit concurrent n'existe), l'estimation des points négatifs et positifs de ce produit à partir d'une évaluation par questionnaire peut constituer un vrai défi (Bangor et al., 2009, 2008). En effet, il n'existe pas, à notre connaissance, de valeurs absolues ou de méthodes permettant l'interprétation directe des questionnaires. On peut tout de même noter les travaux de Bangor, Kortum et Miller (2009, 2008) sur la mise en place de labels d'interprétations pour le questionnaire SUS. Pour cela, ils ont demandé à des participants de compléter

[54] Pour des raisons de confidentialité et de protection de la propriété intellectuelle, cette librairie n'est pas distribuée librement.

à la fois le SUS et un item comprenant plusieurs labels (i.e., « Globalement, je considérerais la facilité d'utilisation (*user-friendliness*) de ce produit comme : la Pire Imaginable, Horrible, Faible, OK, Bien, Excellente, la Meilleure Imaginable » [notre traduction]). Leurs résultats (voir Figure 38) montrent un lien fort entre les adjectifs proposés et les scores au SUS. Ainsi, ces adjectifs peuvent servir comme une forme d'interprétation pour ce questionnaire.

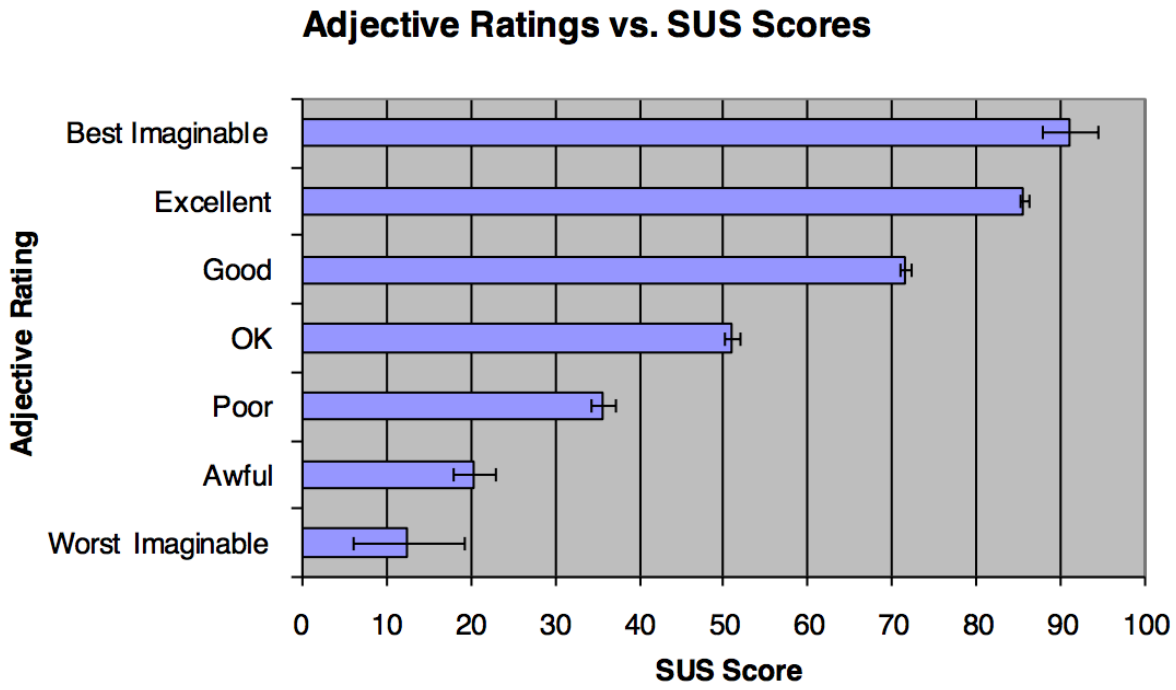


Figure 38. Résultats de l'étude de Bangor, Kortum et Miller (2009, p. 119) : Adjectif choisi en fonction du score au questionnaire SUS

Cependant, cette approche n'est pas applicable à tous les questionnaires. Ainsi, dans une perspective exploratoire, l'usage de régressions logistiques a été testé et semble pouvoir apporter une réponse intéressante.

1.1. Régression logistique

Les régressions logistiques permettent de modéliser l'effet d'une ou plusieurs variables sur une variable binomiale (Hosmer Jr, Lemeshow, & Sturdivant, 2013). Cet outil statistique pourrait permettre, dans le contexte de l'acceptabilité/acceptation des produits technologiques, d'estimer la probabilité d'adopter un produit à partir de l'évaluation d'un ensemble de variables. Après l'estimation des paramètres du modèle (i.e., coefficients bêta), il est possible de réaliser des prédictions sur de nouvelles

données. Par ce biais, il est possible d'estimer la probabilité de survenue d'un événement (e.g., l'adoption d'un produit) en fonction d'un ensemble de paramètres d'entrées (e.g., les dimensions mesurées).

Deux niveaux principaux d'analyses sont envisageables : un niveau d'analyse « produit » et un niveau d'analyse « individuelle ». Le niveau d'analyse « produit » correspondrait à l'estimation de la probabilité qu'un produit soit adopté au regard d'un ensemble d'évaluation par des individus[55]. Ce niveau d'analyse requiert des données diversifiées en termes de produits et de participants pour être le plus fiable possible. Au niveau de l'analyse « individuelle », la probabilité qu'un individu adopte un produit pourrait être estimée.

Afin de tester l'intérêt de cette méthode basée sur des régressions logistiques pour la construction de valeurs seuils, une expérimentation a été réalisée sur des données recueillies avant usage.

1.2. Expérimentation

De nouveaux participants ont été recrutés afin de tester l'approche par régressions logistiques pour l'estimation des valeurs seuils dans les questionnaires.

1.2.1. Méthode de récolte des données

Matériel

Les participants ont évalué l'application de fitness Google Fit. Ce choix devait permettre de recruter des participants n'ayant jamais utilisé ce système mais qui pourraient en avoir l'intention après avoir lu la description. Avant de compléter les échelles, une courte présentation de l'application était proposée aux participants puisqu'ils n'ont jamais utilisé cette application. Cette présentation correspond typiquement à la description proposée par les sociétés sur les magasins d'applications (i.e., AppStore et Google Play). Elle est composée de texte décrivant les fonctions de l'application ainsi que plusieurs captures d'écran de l'application : voir ci-après. « *Nous souhaiterions connaître votre avis sur l'application Google fit décrite ci-dessous.*

[55] Par l'intermédiaire de trois modélisations successives, il est possible d'estimer une probabilité d'adoption ainsi qu'un intervalle de confiance.

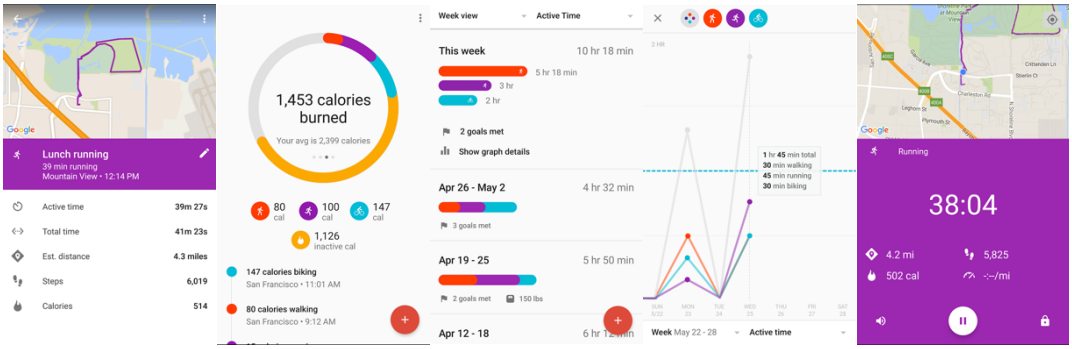


Figure 39. Captures d'écran de l'application Google Fit

Effectuez le suivi de votre activité en toute simplicité. Lorsque vous marchez, courez ou pédalez dans la journée, votre téléphone ou votre montre Android Wear enregistrent automatiquement les données associées à votre activité avec Google Fit ; Obtenez des informations immédiates. Découvrez en temps réel les statistiques liées à vos sessions de course, à vos promenades et à vos sorties à vélo. Google Fit enregistre votre vitesse, votre rythme, l'itinéraire suivi, l'altitude et d'autres informations pour vous aider à rester motivé et à atteindre vos objectifs ; Atteignez vos objectifs de remise en forme. Fixez-vous des objectifs de nombre de pas, de durée, de distance ou de calories brûlées. Recevez des recommandations personnalisées et des encouragements pour atteindre vos objectifs ; Évaluez votre bien-être général. Accédez aux informations d'autres applications de suivi de remise en forme, de nutrition, de sommeil et de poids : Android Wear, Nike+, Runkeeper, Withings et Xiaomi Mi bands ; Utilisez l'application où que vous soyez. Suivez vos progrès sur votre téléphone ou votre tablette, sur le Web, et même sur votre montre Android Wear ».

Participants

146 participants (102 femmes ; moyenne âge = 20.17 ; E-T. âge = 3.04) ont été recrutés au sein de travaux dirigés à l'université Rennes 2. Aucun des participants n'avaient utilisé l'application Google Fit.

Mesures

Les trois échelles (UX anticipée, facteurs affectivo-motivationnels et acceptabilité) présentées dans le Chapitre 4 : Construction des échelles à items uniques pour la mesure avant et après usage ont été utilisées pour évaluer le jugement utilisateur.

La première échelle (échelle UX) permet d'évaluer l'UX anticipée, à savoir l'UX perçue avant usage, à travers 4 dimensions : l'Utilité perçue (« Je trouve que cette application pourrait m'être utile »), l'Utilisabilité perçue (« L'utilisation de cette application semble simple »), la Stimulation (« Cette application me semble novatrice ») et la Confiance perçue (« J'ai confiance dans les informations que cette application pourrait fournir »).

La deuxième échelle (échelle A-M) évalue les facteurs affectifs et motivationnels à travers trois dimensions : la Motivation Intrinsèque (« J'utiliserai cette application parce que je pourrais avoir beaucoup de plaisir à interagir avec cette application), l'Image de Soi (« Je pourrais être perçu(e) positivement par les autres si j'utilisais cette application ») et l'Influence Sociale (« Les personnes de mon entourage qui comptent pour moi pourraient m'encourager à utiliser cette application »).

La troisième échelle évalue les manifestations observables de l'acceptabilité à travers l'Intention d'Usage (« La probabilité que j'utilise cette application est élevée »).

Pour chacun des items, les participants devaient se positionner sur une échelle de Likert en 11 points (de 0 à 10). Afin d'être en mesure de réaliser les estimations des régressions logistiques, une nouvelle question à deux modalités (i.e., oui/non) a été ajoutée au questionnaire : « Je souhaiterais tester cette application ». Cette mesure est utilisée comme la variable binaire à prédire au sein des régressions logistiques.

Procédure

L'application a été présentée avec l'instruction suivante : « Nous voudrions collecter votre opinion sur l'application Google Fit. Afin de vous présenter cette application et ses caractéristiques, vous trouverez ci-dessous une description rapide ». Après avoir lu cette description, les participants devaient compléter les échelles UX et A-M ainsi que les mesures d'acceptabilité via un questionnaire papier.

1.2.2. Résultats et discussion

Plusieurs approches sont envisageables autour de l'utilisation de régressions logistiques pour l'estimation de la probabilité d'adoption. La première approche consiste à estimer un modèle global avec l'ensemble des composantes du jugement utilisateur (e.g., Utilité perçue et Motivation Intrinsèque) comme prédicteur de l'adoption ou du rejet du système. Cette approche peut permettre d'estimer la probabilité globale qu'un individu adopte un système à partir du score qu'il a attribué à

chacune des variables mesurées. Cette méthode a été testée à l'aide de la librairie glm sous R (R Core Team, 2016) à partir des données collectées sur l'application Google Fit. Ainsi, une régression logistique a été estimée en incluant comme prédicteur la Confiance, la Stimulation, l'Utilisabilité, l'Utilité, la Motivation Intrinsèque, l'Image de Soi et l'Influence Sociale. Le modèle converge et montre une influence significative et positive de l'Utilité, la Motivation et l'Influence Sociale sur la variable binaire Intention Comportementale : voir Tableau 24 pour l'ensemble des résultats.

Tableau 24. Résultats de la régression logistique sur l'Intention Comportementale avant usage

	β	E-T	t de Student	p
Utilité	.74	.15	4,93	.00
Utilisabilité	-.08	.13	-.63	.53
Stimulation	-.32	.14	-2,32	.02
Confiance	-.29	.14	-2,10	.04
Influence Sociale	.14	.12	1,15	.25
Image de Soi	-.11	.12	-.99	.32
Motivation	.17	.12	1,34	.18

Bien que cette méthode soit opérationnelle, elle n'est pas celle qui présente le plus d'intérêt. En effet, il est nécessaire d'interroger les participants sur l'ensemble des dimensions pour construire la probabilité d'adoption. Il est vraisemblablement plus efficace de demander directement aux individus s'ils souhaitent ou non adopter le produit plutôt que de passer par une modélisation statistique potentiellement complexe et biaisée.

Une seconde approche consiste à estimer un modèle pour chacune des composantes du jugement utilisateur évaluées. Après estimation des modèles sur chacune des dimensions, cette méthode permet d'évaluer la probabilité d'adoption rattachée à chaque modalité de la dimension mesurée (e.g., voir Tableau 25). Cependant, la fiabilité de cette probabilité est conditionnée à la présence d'un lien significatif entre la composante mesurée et l'intention d'usage. Cette méthode a également été testée sur les données collectées auprès de participants ayant évalué l'application Google Fit. L'ensemble des modèles et résultats associés sont présentés en Annexe 18. Afin d'illustrer le propos, deux modèles sont présentés ci-dessous. Tout

d'abord, une régression logistique avec la variable utilité perçue comme prédicteur de l'intention d'usage a été estimée. Les résultats indiquent un effet significatif de l'utilité perçue (β = .57, p < .001) sur l'intention comportementale et permet d'estimer la probabilité d'adoption associée à chaque modalité de la variable utilité perçue : voir Tableau 25 et Figure 40.

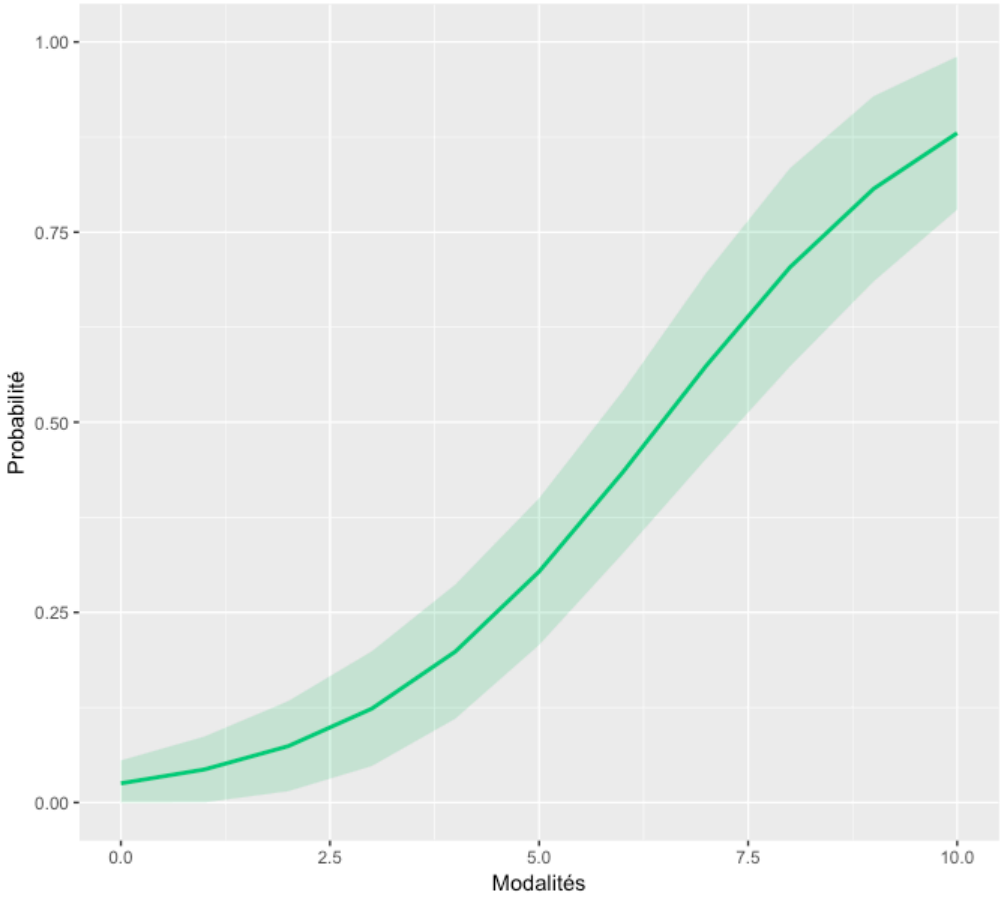


Figure 40. Régression logistique : probabilité d'adoption selon le niveau d'Utilité perçue

Tableau 25. Probabilité d'adoption en fonction du score d'Utilité perçue

Modalité Utilité	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.03	.04	.07	.12	.20	.30	.43	.57	.70	.81	.88

La même approche a été employée avec la variable Utilisabilité perçue. Les résultats montrent une absence d'effet significatif (β = .02, p = .80). Ainsi, des

probabilités d'adoption très similaires apparaissent pour chaque modalité de la variable : voir Tableau 26 et Figure 41.

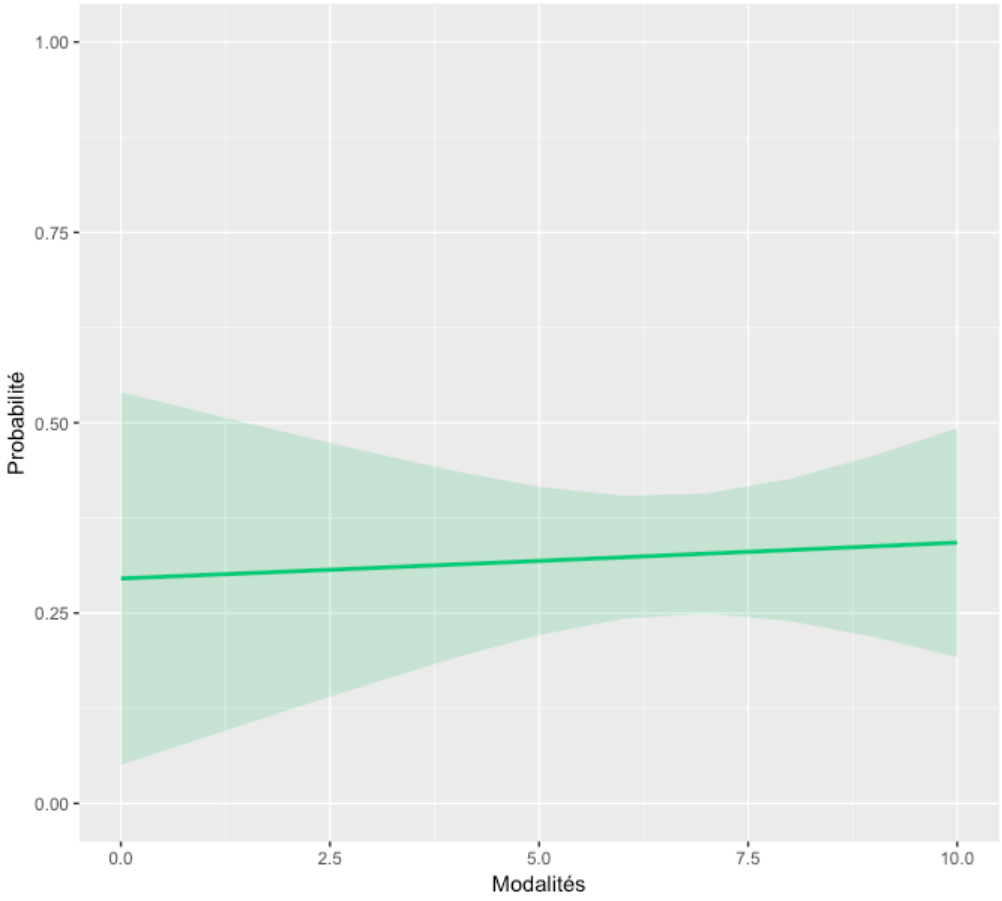


Figure 41. Régression logistique : probabilité d'adoption selon le niveau d'Utilisabilité perçue

Tableau 26. Probabilité d'adoption en fonction du score d'Utilisabilité perçue

Modalité Utilisabilité	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.30	.30	.30	.31	.31	.32	.32	.33	.33	.34	.34

Cette méthode semble apporter des résultats intéressants pour tenter d'estimer des valeurs seuils pour les questionnaires. Plus largement, cette méthode pourrait permettre la détection de points négatifs ou bloquants dans un système. Ces estimations par régressions logistiques seront davantage possibles lorsque les modèles statistiques auront pu être estimés sur de larges échantillons avec une variabilité sur chacune des dimensions. Concrètement, la moyenne des évaluations sur chacune des dimensions sera calculée après avoir réalisé une nouvelle évaluation

d'un système avec un groupe de participants. Ensuite, cette moyenne sera utilisée pour estimer la probabilité d'adoption en se basant sur les modèles de régressions construits pour chacune des dimensions. Par exemple (voir Figure 42), si les participants attribuent en moyenne la valeur 7.5 à la dimension évaluée, la probabilité qu'il adopte le produit est d'environ 62 % en vertu du modèle statistique estimé.

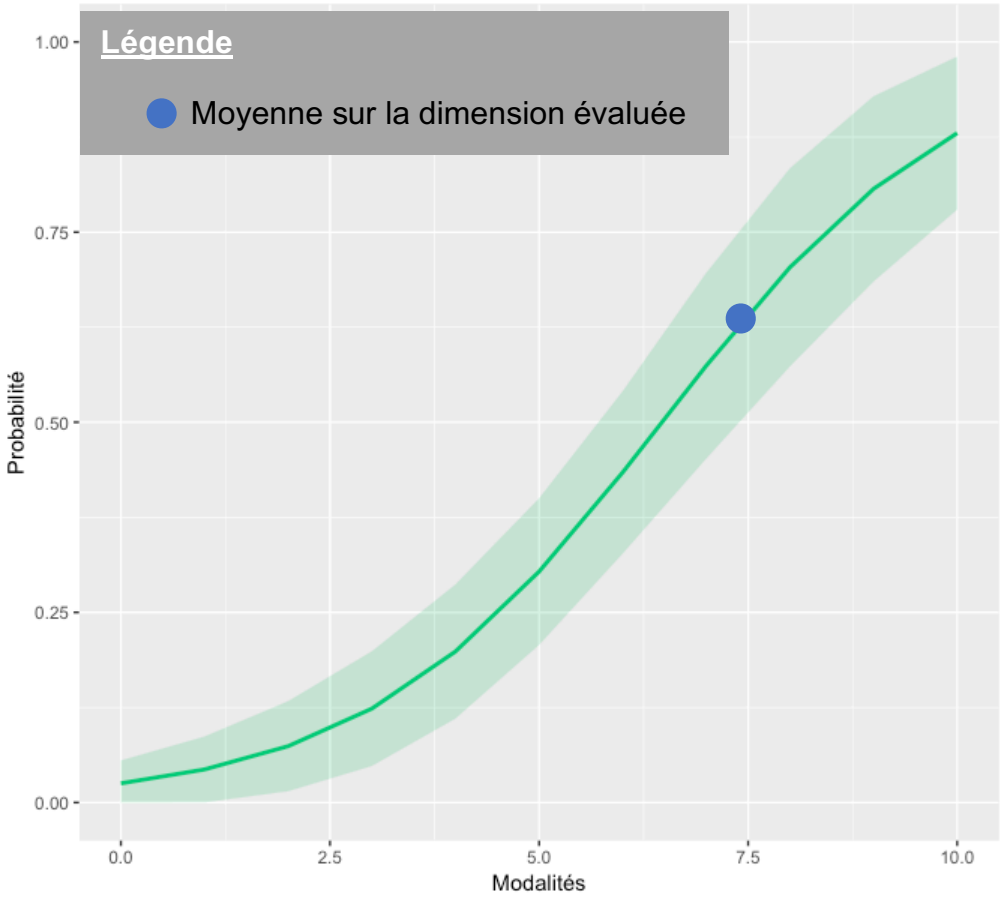


Figure 42. Concept d'estimation de la probabilité d'adoption à partir de la régression logistique et d'une nouvelle valeur[56]

En conclusion, cette approche semble être intéressante pour tenter d'interpréter les questionnaires. Cependant, cette méthode est efficace à condition que la variable évaluée soit un prédicteur significatif. Dans le cas contraire, des résultats comme présentés en Figure 41 sont obtenus. De plus, cette méthode ignore les effets conjoints des variables mesurées puisque les modèles sont estimés en isolant les variables.

[56] Ce graphique n'a qu'une valeur illustrative et ne représente pas une réalité sur des données réelles.

Cette étude présente la seule utilisation au sein de cette thèse des échelles courtes. L'emploi de ces échelles semble réduire les temps de passations et faciliter leur conduite par rapport aux passations avec les échelles à items multiples. Cependant, aucune comparaison formelle n'a pour l'instant été conduite.

Enfin, d'autres analyses ont été conduites sur des données collectées après usage (i.e., acceptation) également à l'aide d'échelles à items uniques. Cependant, des problèmes ont été rencontrés dans les estimations des régressions logistiques. En effet, la variable Intention de poursuivre l'usage présente très peu de variance avec seulement 15 % des participants ayant déclaré l'intention d'arrêter d'utiliser le produit. Cela conduit à un très large déséquilibre entre les groupes. Une méthode de correction du déséquilibre entre les classes (i.e., poursuite ou non poursuite de l'usage), empruntée au domaine de l'apprentissage automatique (He & Garcia, 2009), a été testée. Cependant, la quantité de données initiales n'est pas suffisante et conduit à des problèmes de convergence dans l'estimation des modèles.

Sur la base de l'ensemble des résultats de cette thèse, une librairie R a été conçue et sera présentée dans la partie suivante.

2. Preuve de concept : librairie R

Le langage R et son implémentation GNU R sont de plus en plus utilisés à la fois dans le monde académique et dans le milieu industriel (Muenchen, 2013). De très nombreuses librairies permettent aujourd'hui de réaliser la quasi-totalité des analyses statistiques. De plus, de par son statut *open source*, il est possible de développer ses propres librairies pour réaliser des opérations spécifiques non implémentées dans celles existantes. Ainsi, une librairie R (UXscale) a été développée afin de présenter une preuve de concept et potentiellement faciliter l'exploitation des résultats de cette thèse. Cette dernière comprend principalement une fonction pour l'estimation de scores factoriels pour les échelles à items uniques (voir Chapitre 4 : Construction des échelles à items uniques pour la mesure avant et après usage) en prenant en compte les erreurs de mesure et une fonction pour l'estimation de la probabilité d'adoption basée sur les scores de chaque individu.

Pour l'estimation des scores factoriels à partir des échelles courtes, les fonctions proposées par la librairie lavaan (Rosseel, 2012b) ont été utilisées. En

amont, un modèle factoriel[57] par échelle de mesure (UX, A-M, Acceptabilité) a été estimé en intégrant la variance des erreurs de mesure estimées lors des précédentes études (voir Chapitre 4 : Construction des échelles à items uniques pour la mesure avant et après usage). Au sein de la librairie UXscale, la fonction d'estimation des scores utilise ces modèles afin d'estimer les scores sur de nouvelles données. Le même principe est utilisé pour l'estimation de probabilité d'adoption. Des modèles de régressions logistiques ont été estimés à l'aide de la fonction glm de la librairie standard. Ces modèles sont utilisés par la librairie UXscale pour émettre des prédictions sur de nouvelles données. D'autres briques pour faciliter l'usage ont été ajoutées notamment en s'appuyant sur les librairies fmsb (Nakazawa, 2017) et ggplot2 (Wickham, 2009) : estimations de statistiques descriptives, créations de graphes de points (*dotchart*), de graphes radars et d'histogrammes.

En résumé, l'environnement et le langage de programmation R offrent de nombreuses perspectives en permettant de concevoir de nouvelles librairies adaptées à des besoins spécifiques. Cependant, afin d'optimiser le temps de développement, l'approche classique consiste à utiliser d'autres librairies reconnues auxquelles sont ajoutées de nouvelles fonctions. Ce fonctionnement peut conduire à certains problèmes. En effet, de par la dépendance à des librairies externes, des modifications dans ces dernières peuvent avoir des effets secondaires sur la librairie développée. La jeunesse de l'environnement R et son statut collaboratif conduisent à de nombreuses corrections et améliorations sur les différentes librairies. Il est ainsi nécessaire de maintenir la librairie régulièrement à jour pour s'adapter à d'éventuelles modifications dans d'autres librairies.

En conclusion, cette librairie pourrait servir dans un futur proche, au terme d'un développement logiciel plus important, d'outil d'interprétation simple et rapide pour les praticiens. L'idéal serait de pouvoir proposer un logiciel interprétant directement les données brutes obtenues à travers le questionnaire.

[57] Trois modèles ont été construits : un modèle avec la valeur moyenne de chaque erreur de mesure, un modèle avec la valeur de la borne inférieure de chaque erreur de mesure et un modèle avec la valeur de la borne supérieure de chaque erreur de mesure. Cette méthode permet d'obtenir l'intervalle de confiance dans lequel chaque valeur devrait se situer.

Chapitre 6 : Discussion Générale

Cette thèse avait pour objectif de développer un nouvel outil de mesure du jugement utilisateur envers des produits technologiques et d'analyser les liens entre le jugement utilisateur et ses composantes. Depuis l'apparition des premiers systèmes technologiques, la compréhension des origines de l'adoption de ces systèmes par les utilisateurs est un sujet important pour le domaine des Interactions Homme-Machine que cela soit pour les chercheurs ou les praticiens. Ce long et complexe processus d'adoption repose sur de nombreuses composantes du jugement utilisateur. Au sein de cette thèse, ces éléments ont été explorés à différents moments du processus d'adoption ou de rejet du système. Ainsi, les principaux résultats, les limites identifiées, l'impact de nos travaux à la fois pour la recherche et les praticiens ainsi que les perspectives de recherche seront exposés.

1. Acceptabilité, UX et variables affectivo-motivationnelles

Préalablement à la synthèse des principaux résultats, un rappel des principaux travaux théoriques sur lesquels s'appuient cette thèse sera proposé.

1.1. Résumé des fondements théoriques

1.1.1. Modèles théoriques

Ces travaux de thèse ont comme point de départ le constat suivant : plusieurs approches théoriques sur le jugement utilisateur envers des produits technologiques coexistent mais sont rarement mises en lien. Ces différents champs théoriques partagent la volonté de mettre en lumière les dimensions liées au jugement d'un utilisateur envers un système mais avec des finalités légèrement différentes. En effet, les travaux sur l'acceptation visent à identifier les facteurs qui conduisent à l'adoption ou au rejet d'un système (e.g., Balog & Pribeanu, 2010 ; Bruner & Kumar, 2005). Ils s'appuient majoritairement sur les modèles théoriques TAM (F. D. Davis, 1989) et UTAUT (Venkatesh et al., 2003 ; Venkatesh, Thong, & Xu, 2012b). Ces derniers mettent en avant l'effet des variables liées aux qualités perçues d'un produit sur

l'intention comportementale. En ce qui concerne les travaux sur l'UX, ils visent à mettre en lumière les facteurs du jugement global envers un système technologique. Plus récemment, ces modèles ont évolué pour inclure d'autres variables que la perception de qualités fonctionnelles du système. Le modèle d'Hassenzahl (2004) y fait figure de référence en distinguant la perception de qualités instrumentales (i.e., utilité perçue) et la perception des qualités non instrumentales (e.g., identification) comme prédicteur du jugement global. Parallèlement, d'autres variables, dites affectivo-motivationnelles (e.g., motivation intrinsèque ; F. D. Davis et al., 1992), ont été incluses dans ces modèles afin d'appréhender au mieux les divers éléments composant le jugement utilisateur. L'intégration de ces variables vise également à dépasser les critiques relatives à la vision trop rationnelle du jugement (i.e., l'adoption d'un système serait le résultat de l'évaluation rationnelle des qualités de celui-ci) sur laquelle reposent certains modèles (Brangier et al., 2010). Ainsi, intégrer au sein d'un modèle unifié ces différents éléments théoriques représentent une approche intéressante et rarement mise en application (van Schaik & Ling, 2011). En ce sens, cette démarche permet d'améliorer la compréhension de l'adoption des systèmes technologiques, jugement évoluant, quant à lui, au cours du temps.

1.1.2. Temporalité dans le jugement

L'adoption et le jugement d'un système technologique recouvrent différentes temporalités (Bobillier-Chaumon & Dubois, 2010). En effet, la formation du jugement commence avant l'utilisation du système avec les représentations préalables et évolue jusqu'à l'adoption ou le rejet du produit à long terme. Les différences au cours du temps et notamment entre l'avant et l'après usage ont été mises en lumière au sein des conceptions théoriques de l'UX (ISO, 2010) et de l'acceptation (Schade & Schlag, 2003). Cependant, le jugement avant usage a rarement fait l'objet de travaux expérimentaux (Šumak & Šorgo, 2016 ; ogasara et al., 2011). S'intéresser à cette première étape dans la construction de l'intention d'usage semble un élément important pour compléter les connaissances sur l'adoption des produits technologiques. Ainsi, les études ont été conduites à la fois sur le jugement avant usage et après usage.

1.2. Principaux résultats

Au cours de la première étude, concernant l'avant usage, il a été demandé aux participants d'évaluer l'application Sleep Better à partir de la description proposée par ses concepteurs sur les magasins d'applications. Il a été montré un effet positif et significatif de l'utilité perçue et de la motivation intrinsèque sur l'intention d'usage. Ce résultat a été répliqué dans une seconde étude où un nouveau groupe de participants a évalué l'application mobile Google Fit.

Une troisième étude a été réalisée sur le jugement après usage. Il a été demandé à des utilisateurs d'évaluer l'application mobile Waze, après l'avoir utilisée. Les données obtenues ont permis de mettre en avant une influence significative et positive de l'utilité perçue ainsi qu'un effet négatif significatif des émotions négatives sur l'intention d'usage.

Enfin, une dernière étude a été menée pour estimer les éventuels changements entre le jugement avant et après usage envers une application mobile chez le même groupe de participants. Les résultats témoignent d'une influence négative et significative de l'usage sur l'intention comportementale, l'utilité, l'utilisabilité, la confiance, l'image de soi et l'influence sociale. De plus, en se basant sur le cadre théorique de l'EDT (*Expectations-Disconfirmations Theory* ; Brown, Venkatesh, & Goyal, 2012), il a été montré que les attentes (i.e., jugement avant usage) en termes d'utilité, d'utilisabilité et de stimulation influenceraient significativement l'intention comportementale. Deuxièmement, les infirmations (i.e., différence entre les perceptions initiales et la réalité) sur l'utilité, l'utilisabilité, la stimulation et la confiance seraient des prédicteurs significatifs de l'intention d'usage. Enfin, parmi les facteurs affectivo-motivationnels, la motivation intrinsèque semble être le seul facteur réellement prédicteur de l'intention comportementale. En d'autres termes, les perceptions initiales des qualités ainsi que les différences entre ces perceptions et la perception après usage auraient une influence sur l'intention de poursuivre l'usage d'un produit. Ainsi, il semble intéressant de prendre en compte l'acceptabilité en évaluant avant usage la perception des qualités fonctionnelles et non fonctionnelles ainsi que des facteurs affectivo-motivationnels pour appréhender au mieux la poursuite de l'usage au cours du temps.

En résumé, les résultats témoignent d'un rôle majeur de l'utilité perçue que cela soit avant ou après usage, effet répliqué au sein de toutes les études. Ce résultat est

cohérent avec les méta-analyses existantes sur l'acceptation (e.g., Dwivedi et al., 2011 ; Khechine et al., 2016). L'influence des dimensions affectivo-motivationnelles varie en fonction des temporalités d'évaluations. La motivation intrinsèque aurait un effet uniquement avant usage alors que les émotions négatives présenteraient un effet uniquement après usage. À l'inverse, les qualités non fonctionnelles des modèles de l'UX ne semblent pas avoir une influence sur l'intention comportementale dans les différentes études réalisées.

Au sein du chapitre théorique (voir en page 67), un modèle unifié avait été proposé à partir des précédents travaux de la littérature. À la lumière des résultats de cette thèse, ce modèle peut être mis à jour : voir Figure 43.

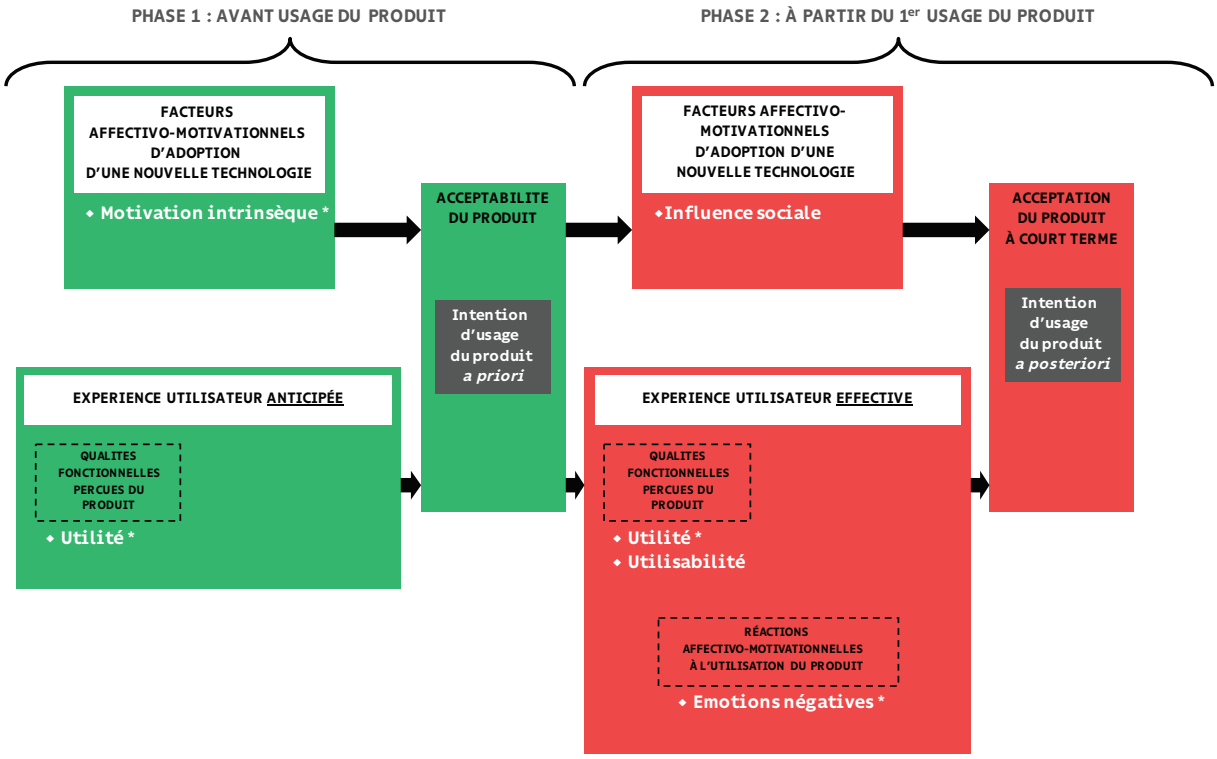


Figure 43. Modèle intégratif et dynamique de l'acceptabilité/acceptation d'une nouvelle technologie mis à jour au regard des résultats des études[58]

Les effets significatifs de certaines variables, pourtant bien étayés dans la littérature (e.g., utilisabilité ; Khechine et al., 2016 ; King & He, 2006), n'ont pas été reproduits. Cette absence d'effet pourrait s'expliquer par le fait que les résultats significatifs ont tendance à être davantage publiés dans la littérature (Song et al.,

[58] * = variable ayant un effet significatif.

2010). Ce biais pourrait ainsi conduire à surestimer à partir des résultats précédents les effets de certaines variables. Cela pourrait être d'autant accentué par les qualités des systèmes qui ont été sélectionnés pour être évalués.

2. Échelles de mesure

Afin d'évaluer le modèle unifié du jugement utilisateur et de répondre aux objectifs du projet UXPPE, une partie importante des travaux a été dédiée à la création d'un outil de mesure. Le choix s'est porté sur la création d'un questionnaire comme dans de nombreuses précédentes études (Berkman & Karahoca, 2016). Les questionnaires permettent en effet de collecter facilement une grande quantité de données, avantage pour les praticiens et les chercheurs, afin de tester quantitativement les modèles théoriques. Les travaux autour de la création des questionnaires se sont déroulés en deux étapes. Une première version des échelles a été construite sur la base des travaux sur les variables affectivo-motivationnelles et des recherches liées aux champs théoriques de l'UX et de l'acceptation. Composée de plusieurs items par dimension évaluée pour l'avant et l'après usage, cette première version a été testée au cours des différentes études présentées précédemment. Les analyses factorielles exploratoires, confirmatoires ainsi que l'analyse de l'invariance[59] ont mis en avant des qualités psychométriques satisfaisantes (avec principalement des structures factorielles appropriées). Cependant, d'autres analyses mériteraient d'être exécutées. Il serait notamment nécessaire d'accomplir de nouvelles collectes de données en proposant à la fois les échelles développées dans cette thèse ainsi que d'autres outils présents dans la littérature (e.g., Attrackdiff ; Hassenzahl, 2004) afin de s'assurer de la validité concourante (i.e., corrélation entre les scores du nouvel outil et les scores de l'outil qu'il remplace ; Campbell & Fiske, 1959). La validité test-retest (i.e., le questionnaire est complété à deux moments différents) serait également à évaluer afin de s'assurer de la fidélité des questionnaires (McCrae, Kurtz, amagata, & Terracciano, 2011). Ces approches permettent de s'assurer au mieux de la validité

[59] Les analyses sur l'invariance factorielle n'ont pu être réalisées que sur le jugement avant usage où deux échantillons indépendants ont pu être recrutés. En ce qui concerne le jugement après usage, un seul échantillon était utilisable et par conséquent l'invariance factorielle n'a pas pu être testée.

des échelles construites. Cependant, elles n'ont pas été mises en place face à la difficulté de recruter suffisamment de participants. En effet, les participants ont eu tendance à abandonner lors de la complétion des questionnaires, en particulier lorsque celui-ci était long. Ainsi, augmenter la taille des questionnaires en rajoutant d'autres échelles de la littérature aurait assurément augmenté les abandons.

La seconde partie des travaux sur les outils d'évaluation a été centrée sur la réduction de la longueur des questionnaires à travers la construction d'échelles à items uniques. Dans le même temps, une approche a été explorée pour limiter les problèmes liés à l'utilisation de questionnaires courts en prenant en compte les erreurs de mesure. Enfin, une preuve de concept a été proposée pour estimer des valeurs de seuils au sein des questionnaires afin de faciliter leur interprétation.

Les travaux réalisés sur les questionnaires présentent des résultats tout à fait satisfaisants et ont permis l'élaboration d'un outil pouvant être rapidement transférable vers les praticiens. Cependant, l'outil de mesure développé n'échappe pas aux critiques classiques des questionnaires. Dans le cadre de l'évaluation de systèmes technologiques, les données collectées à l'aide d'échelles ne fournissent pas toujours beaucoup d'informations sur les éléments qu'il pourrait être nécessaire de faire évoluer dans le produit (Lallemand & Gronier, 2015). Or, les données collectées lors des tests utilisateurs, notamment dans une démarche de CCU, doivent servir à améliorer le système afin de répondre au mieux aux besoins des utilisateurs. De plus, il est nécessaire d'utiliser les questionnaires standardisés dans leur configuration originale pour s'assurer de la généralisation des données. Or, certaines situations d'évaluation peuvent nécessiter de modifier les outils de mesure fournis par la littérature. Bien que ce besoin soit totalement compréhensible, ces outils « faits maison » peuvent poser des problèmes sur le plan psychométrique et conduire à des évaluations de mauvaise qualité (Morse, 2009). D'autres pistes d'améliorations sont également à signaler et seront présentées dans la partie suivante.

3. Pistes d'améliorations et futurs travaux

Plusieurs limites et pistes d'améliorations peuvent être envisagées pour améliorer la compréhension des facteurs du jugement utilisateur et guider de futurs travaux.

3.1. Produits évalués et conditions expérimentales

Pour étudier le jugement utilisateur et ses composantes, plusieurs études ont été réalisées, au cours desquelles les participants ont évalué plusieurs applications mobiles. La sélection de ces applications comme matériel expérimental résulte d'une longue réflexion prenant en compte plusieurs variables. Tout d'abord, comme présenté précédemment, il a été choisi de conduire des études en ligne. Ainsi, il était nécessaire que les participants puissent accéder au produit gratuitement et à distance. De plus, les recherches devaient inclure l'évolution du jugement utilisateur depuis les représentations avant l'usage à l'usage prolongé. Il était donc nécessaire de trouver des situations où les produits devaient être évalués avant usage. Le choix s'est alors porté sur des applications mobiles répondant à ces diverses exigences.

Bien que le matériel expérimental possède de nombreux avantages pour la réalisation de passations, il peut faire l'objet de plusieurs critiques. Premièrement, les expérimentations n'ayant eu lieu que sur des applications mobiles gratuites, la généralisation des résultats est à nuancer. En effet, une large partie des produits technologiques est payante et parfois à des sommes élevées. Il est envisageable que l'effet de certaines variables puisse être différent lorsque les individus doivent choisir des produits payants. Venkatesh et ses collaborateurs (2012b) ont, par exemple, montré une influence de la valeur de prix (i.e., « le compromis cognitif des consommateurs entre les avantages perçus des applications et le coût monétaire pour les utiliser » [notre traduction] ; Dodds et al., 1991) sur l'intention d'usage. Cependant, il est très complexe de conduire ce type d'étude sans le concours d'un industriel. Deuxièmement, les études ont été uniquement menées sur des applications mobiles. Des différences pourraient apparaître lors de l'étude d'autres types de produits. Troisièmement, les qualités des systèmes n'ont pas été manipulées au sein des études. Cette méthode, classique en psychologie expérimentale, permet de tester des hypothèses précises sur certaines variables. Cette approche n'a pas été mise en place pour plusieurs raisons. Faire varier les paramètres des produits peut être une opération complexe avec des résultats incertains. En effet, ce type de modification expérimentale peut mener à des interfaces très artificielles pouvant apparaître comme non réalistes auprès des participants. De plus, la manipulation de certaines variables pourrait influencer l'évaluation d'autres variables (e.g., changement de la perception de l'utilisabilité lors de la modification expérimentale de l'esthétisme ; Tractinsky et al.,

2000) ce qui peut compliquer l'isolation des variables les unes par rapport aux autres (Tuch, Roth, Hornbæk, Opwis, & Bargas-Avila, 2012). Enfin, manipuler les composantes d'un produit nécessite d'avoir la possibilité de travailler avec une équipe de développement logiciel, ce qui n'était pas le cas dans le cadre de ce projet de thèse. Enfin, d'autres travaux pourraient être réalisés en faisant varier les descriptions fournies aux participants durant les évaluations avant usage. Cela permettrait d'estimer plus finement les éléments influençant le jugement des individus avant qu'ils n'aient pu tester le système.

3.2. Échantillons et méthodes de recueils des données

3.2.1. Echantillon

Les échantillons recrutés sont majoritairement composés de participants jeunes (âge moyen = 28 ans). Ainsi, des différences peuvent être envisagées dans le jugement en fonction de l'âge des utilisateurs (e.g., Czaja et al., 2006 ; Venkatesh et al., 2003). De plus, une surreprésentation des femmes apparaît au sein des différents échantillons de participants. Cela pourrait également avoir une influence sur les résultats. De précédentes études ont montré des effets différents en fonction du sexe (e.g., Venkatesh et al., 2003, 2012b).

3.2.2. Études en ligne

Des modèles structuraux ont été utilisés pour construire l'outil de mesure et tester les hypothèses construites à partir de la littérature. En effet, les SEMs offrent un cadre d'analyse unifié (Hoyle, 2012) pour évaluer à la fois le modèle de mesure (i.e., liens entre les mesures et les variables latentes) et le modèle structurel (i.e., liens entre les variables du modèle). Cependant, l'estimation de ces modèles requiert la participation de nombreux individus (Kline, 2010). Ce type de recrutement est une opération fastidieuse et coûteuse sur le plan temporel et financier. Pour tenter de faciliter la collecte de ces données, la majeure partie du recrutement des participants

dans cette thèse a eu lieu en ligne et notamment sur les réseaux sociaux[60]. Cette approche offre de nombreux avantages et, par conséquent, se développe de plus en plus (Bartneck, Duenser, Moltchanova, & Zawieska, 2015 ; Gosling, Vazire, Srivastava, & John, 2004). De ce fait, il devient possible d'atteindre des populations plus importantes et diversifiées que lors des recrutements réalisés dans les universités (Bartneck et al., 2015 ; Krantz & Dalal, 2000), comme couramment pratiqués pour les études en psychologie. D'autre part, l'anonymat procuré par internet permettrait aux participants de répondre aux questions même lorsqu'elles évoquent des sujets sensibles comme la consommation de drogue (G. S. Ryan, 2013). Les récoltes de données informatisées offriraient également la possibilité de réduire les temps d'analyses. Elles diminueraient les erreurs liées à la saisie de données et permettraient de modifier facilement le questionnaire (Wyatt, 2000) en introduisant, par exemple, la randomisation de l'ordre de présentation des questions afin d'éviter les effets d'ordre (Benton & Daly, 1991).

Cependant, d'autres chercheurs s'interrogent sur la qualité des données obtenues, via ce mode de recueil. Les participants seraient moins impliqués dans les études car ils seraient moins motivés. Par conséquent, ils auraient tendance à répondre davantage de manière aléatoire. Néanmoins, selon Gosling et collaborateurs (2004), il n'y aurait pas de preuve d'une plus faible motivation (affectant les données) chez les personnes prenant part à une étude en ligne que chez ceux participant à une étude papier-crayon. D'autres critiques sont également émises envers les études en ligne. Celles-ci pointent les difficultés de filtrage des participants (e.g., les participants pourraient remplir plusieurs fois le questionnaire) et l'incapacité, pour l'expérimentateur, de contrôler les passations (e.g., impossibilité de contrôler la diffusion du questionnaire et donc de calculer un taux de réponse ; Wyatt, 2000). Pour limiter ces biais, diverses méthodes ont été employées durant les passations de cette thèse : l'introduction d'un code unique par participant et/ou le blocage des doubles enregistrements (en cas de mauvais clic ou clic intempestif). Bien que des biais potentiels liés à ce mode de recueil puissent légèrement affecter les données, les études en ligne offrent la possibilité d'accéder à des tailles d'échantillon difficilement

[60] Plusieurs approches existent pour le recrutement sur les réseaux sociaux comme le paiement de publicité ou encore le partage sur des groupes Facebook (Yuan, Bare, Johnson, & Saberi, 2014).

envisageables autrement. Néanmoins, il pourrait être intéressant de comparer les résultats obtenus, au sein de cette thèse, avec des échelles présentées en laboratoire.

Enfin, une plate-forme d'enregistrement des données a été créée spécifiquement pour la collecte des données. Cette dernière, adaptée à la fois à un usage mobile et ordinateur, devait permettre de répondre aux recommandations de la CNIL (Commission Nationale de l'Informatique et des Libertés) et à la confidentialité sur les données réclamées par l'IRT b<>com. Cet outil offre également la possibilité d'enregistrer les abandons des participants, randomiser la présentation des questions et peut permettre un déploiement à plus large échelle du questionnaire en ligne. Pour collecter de nombreuses données et faciliter un déploiement auprès des praticiens, cette plate-forme pourrait être étendue et améliorée. Par exemple, il pourrait être possible de faire profiter gratuitement de la plate-forme incluant le questionnaire en échange de la mise à disposition des données par les expérimentateurs. Cette méthode faciliterait la distribution du questionnaire en même temps qu'elle donnerait la possibilité d'affiner les modèles et tester de nouvelles hypothèses.

3.3. Modèles proposés

L'état de l'art des divers modèles et conceptions théoriques du jugement utilisateur a permis de dégager des variables et modèles apparaissant comme pertinents au regard de ces travaux de thèse. Cependant, les contraintes imposées par le projet UXPPE ont eu des influences importantes sur les travaux de cette thèse. En effet, il a été choisi de travailler sur des modèles facilement utilisables pour les praticiens. Ainsi, les modèles testés ne supposent que des effets directs des variables sur l'intention comportementale. Or, des modèles et liens plus complexes entre les variables sont à supposer au regard de la littérature (e.g., effet indirect de l'utilisabilité perçue via l'utilité perçue ; King & He, 2006). Toutefois, la mise en place de ces modèles, bien que très intéressante sur le plan théorique (voir exemple en Annexe 19 sur des données collectées au cours de la thèse) pose des problèmes dans leur application. En effet, il est nécessaire pour les praticiens de pouvoir isoler les variables sur lesquelles ils pourront influer en modifiant le système. Ils pourront ainsi en améliorer le jugement et potentiellement faciliter l'adoption par les utilisateurs. Or, des modèles plus complexes, incluant, par exemple, des effets indirects, peuvent rendre difficile les interprétations et les recommandations pour d'éventuelles modifications

des systèmes étudiés. Ainsi, des modèles plus simples semblaient mieux répondre aux besoins des praticiens même s'ils peuvent présenter une vision davantage tronquée de la réalité.

Enfin, il s'avèrerait intéressant de réaliser une méta-analyse sur les modèles de l'UX : me-CUE (Thüring & Mahlke, 2007b) et Hassenzahl (2005b). En effet, dans le domaine de l'acceptabilité, les principaux modèles, TAM (F. D. Davis, 1989) et UTAUT (Venkatesh et al., 2003), ont largement été utilisés dans de nouvelles études. Ces recherches ont fait l'objet de plusieurs méta-analyses (e.g., Khechine et al., 2016 ; Ma & Liu, 2004) qui ont permis de confirmer les liens proposés dans ces modèles. Réaliser le même exercice sur les modèles de l'UX permettrait d'obtenir une meilleure vue d'ensemble des travaux et des résultats du domaine. De plus, comme présenté précédemment, ces deux champs théoriques ne sont que très rarement mis en relation à l'exception des travaux de van Schaik et collaborateurs (Aranyi & van Schaik, 2015a, 2015b ; van Schaik & Ling, 2011). Or, ces modèles semblent complémentaires et apporter de nouvelles informations sur le jugement utilisateur. De plus, les outils de mesure existants, notamment de l'UX, mériteraient de nouvelles analyses pour s'assurer de leurs qualités psychométriques (e.g., Lallemand, Koenig, et al., 2015 ; J. R. Lewis, 2013).

3.4. Mesure du jugement utilisateur et de l'adoption

La mesure de l'usage réel pourrait également être améliorée dans de futurs travaux. En effet, comme dans la majorité des recherches précédentes (Sezgin & ıldırım, 2016), l'évaluation de l'intention comportementale a servi de mesure à l'adoption d'un système (e.g., L. Chen et al., 2012). L'usage réel qui en découle est très rarement exploré, ce qui s'explique pragmatiquement par les difficultés à le mesurer. Dans la littérature, lorsqu'une évaluation de l'usage réel est réalisée, elle repose généralement sur du déclaratif (e.g., Erasmus, Rothmann, & van Eeden, 2015 ; Venkatesh et al., 2012b) et très rarement sur une mesure objective de l'usage. La mesure objective nécessite, en effet, de pouvoir accéder directement aux données d'usage du système. Au-delà des contraintes liées à la protection des données, un partenariat direct avec une entreprise développant un produit (e.g., application mobile) est généralement requis, ce qui n'a pas été possible de mettre en place durant cette thèse.

3.5. Temporalité

Plusieurs temporalités du jugement utilisateur ont été étudiées au cours de cette thèse. Concernant le jugement avant usage, à notre connaissance, très peu de travaux dans la littérature ont été conduits sur l'acceptabilité (Šumak & Šorgo, 2016) ou l'UX anticipée (ogasara et al., 2011) et ce, encore moins dans une approche unifiée. La majorité des travaux est, en effet, centrée sur le jugement après usage. Deux raisons principales peuvent expliquer cela. Premièrement, les industriels souhaitent généralement connaître le jugement des utilisateurs sur un produit qu'ils ont potentiellement développé dans une démarche de CCU (norme ISO 9241-210 ; 2010). Cependant, dans la majorité des développements de produits technologiques, la démarche CCU n'est pas toujours appliquée et les évaluations avec les utilisateurs commencent souvent lorsque les premiers prototypes ont été conçus ou parfois même sur les produits quasi finis. Deuxièmement, il est très difficile de recruter un groupe de participants de taille suffisamment importante pour évaluer un produit avant usage dans des conditions écologiques (i.e., produit déjà existant mais que les participants n'ont pas encore testé). Dans le cadre de cette thèse, le matériel expérimental a été choisi de façon à pouvoir potentiellement être testé auprès des participants après l'évaluation avant usage. Ainsi, il était nécessaire de trouver un produit gratuit et facilement accessible que les participants n'avaient jamais testé. Cela était d'autant plus difficile que lorsque le produit est gratuit et de bonne qualité, sa popularité a tendance à rapidement augmenter.

Concernant le jugement après usage, un seul échantillon a pu être exploité car les autres présentaient un nombre trop réduit de participants. Les résultats montrent des différences importantes par rapport à la littérature. En effet, l'utilisabilité ne présente pas d'effet significatif alors que cette variable est bien étayée dans la littérature (e.g., Khechine et al., 2016 ; King & He, 2006). Afin de s'assurer des résultats obtenus dans le cadre du modèle unifié de l'acceptation, de l'UX et des variables affectivo-motivationnelles, il semble nécessaire de tenter de répliquer cette étude. De plus, selon de précédentes études, le jugement devrait évoluer avec l'usage (e.g., Harbich & Hassenzahl, 2016 ; Wimmer et al., 2010). Dans le cadre de ces travaux, seulement un moment d'observation après usage a finalement pu être effectué. Afin de mieux percevoir les changements au cours du temps sur les variables

mesurées, conduire une étude longitudinale incluant notamment des mesures objectives de l'usage réel apparaît comme une perspective intéressante.

Enfin, concernant les différences entre le jugement avant et après usage, les processus d'évaluation diffèrent dans les fonctions cognitives auxquelles ils font appel. En effet, lors du jugement avant usage, il est demandé aux participants de réaliser un jugement prospectif. Les participants doivent ainsi imaginer le système et ses différentes qualités à partir des informations qu'ils ont à leur disposition (e.g., description du système, publicité, avis de l'entourage) pour construire leurs perceptions (Bhattacherjee & Premkumar, 2004). Le postulat effectué est que les participants sont en mesure d'imaginer le système ainsi que l'interaction qu'ils auront avec celui-ci (Pakanen, 2015). Dans le cadre des études de cette thèse, il est possible que certaines qualités telles que l'utilisabilité aient été plus difficiles à évaluer à partir des descriptions proposées. L'évaluation reposerait alors sur des informations antérieures et/ou connexes (i.e., précédentes expériences avec des systèmes similaires ou perception des capacités d'utilisation) ou correspondrait à des réponses aléatoires. Dans le cadre du jugement après usage, les participants doivent émettre un jugement rétrospectif. Ce type d'évaluation peut amener à surestimer certains événements en conséquence de certains biais cognitifs, comme l'effet de récence et de primauté (i.e., les événements les plus anciens ainsi que les plus récents auront tendance à être mieux rappelés) ou la *peak-end rule* (les derniers éléments et ceux ayant suscités les plus fortes intensités émotionnelles auront une influence plus élevée ; Geng, Chen, Lam, & Zheng, 2013).

Conclusion

Les travaux de cette thèse visaient à construire un modèle unifiant le champ théorique de l'acceptabilité et celui de l'UX, dans le but de proposer un outil de mesure le plus complet possible du jugement utilisateur. Ainsi, l'originalité de ces travaux se situe dans l'intégration de ces deux champs théoriques avec l'inclusion des variables affectivo-motivationnelles mais également dans la conduite d'études à la fois sur l'avant et l'après usage.

En résumé, plusieurs principaux résultats peuvent être mis en lumière. Sur le plan théorique, les résultats soulignent l'influence majeure de l'utilité perçue à travers les temps d'évaluations. Les variables affectivo-motivationnelles, quant à elles, présentent des effets différents en fonction de la temporalité du jugement. Enfin, les études ne montrent aucune influence significative des qualités d'UX non fonctionnelles. Sur le plan applicatif, un outil de mesure, via l'élaboration d'un questionnaire comprenant une évaluation des facteurs d'acceptabilité/acceptation, d'UX et affectivo-motivationnels avant et après usage, a été construit et testé. Cet outil, disponible en version à items multiples et à items uniques, pourra servir à la conduite de nouvelles études afin de confirmer le modèle unifié proposé ou évaluer des prototypes fonctionnels auprès d'utilisateurs dans le cadre d'une CCU. En conclusion, il semble nécessaire de continuer le développement d'outils valides, simples d'utilisation et faciles d'interprétations, à disposition des praticiens, unifiant les différentes approches et champs théoriques.

Références

Abdel-Khalek, A. M. (2006). Measuring Happiness with a single-item scale. *Social Behavior and Personality: an international journal*, *34*(2), 139‑150. https://doi.org/10.2224/sbp.2006.34.2.139

Abdi, H. (2003). Factor Rotations in Factor Analyses. Dans M. Lewis-Beck, A. Bryman, & T. F. Liao (Éds), *Encyclopedia of Social Science Research Methods*. Thousand Oaks, Calif : SAGE Publications Ltd.

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433‑459. https://doi.org/10.1002/wics.101

Adell, E., Varhelyi, A., & Nilsson, L. (2014). The Definition of Acceptance and Acceptability. Dans M. Regan (Éd.), (pp. 11–22). (S.l.) : Ashgate.

Agarwal, R., & Karahanna, E. (2000). Time Flies When ou're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. *MIS Quarterly*, *24*(4), 665‑694. https://doi.org/10.2307/3250951

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179‑211. https://doi.org/10.1016/0749-5978(91)90020-T

Alagumalai, S., & Curtis, D. D. (2005). Classical Test Theory. Dans D. R. Maclean, R. Watanabe, R. Baker, D. Boediono, P. . C. Cheng, D. W. Duncan, … N. Hungi (Éds), *Applied Rasch Measurement: A Book of Exemplars* (pp. 1‑14). (S.l.) : Springer Netherlands. https://doi.org/10.1007/1-4020-3076-2_1

Alavifar, A., Karimimalayer, M., & Anuar, P. M. K. (2012). Structural equation modeling VS multiple regression. *IRACST – Engineering Science and Technology: An International Journal (ESTIJ)*, *2*(2), 326‑329.

Allison, P. D. (2003). Missing Data Techniques for Structural Equation Modeling. *Journal of Abnormal Psychology*, *112*(4), 545‑557. https://doi.org/10.1037/0021-843X.112.4.545

Amaro, S., Abrantes, J. L., & Seabra, C. (2015). Comparing CB-SEM and PLS-SEM results: an empirical example. University of Twente. https://doi.org/10.3990/2.357

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411‑423. https://doi.org/10.1037/0033-2909.103.3.411

Aranyi, G., & van Schaik, P. (2015a). Modeling user experience with news websites. *Journal of the Association for Information Science and Technology*, 2471–2493. https://doi.org/10.1002/asi.23348

Aranyi, G., & van Schaik, P. (2015b). Testing a model of user-experience with news websites. *Journal of the Association for Information Science and Technology*, 1555–1575. https://doi.org/10.1002/asi.23462

Arhippainen, L., & Tähti, M. (2003). Empirical evaluation of user experience in two adaptive mobile application prototypes. Communication présentée au MUM, Norrköping, Sweden.

Attfield, S., Kazai, G., Lalmas, M., Piwowarski, B., Attfield, S., Kazai, G., … Piwowarski, B. (2011). Towards a science of user engagement. Communication présentée au UMWA 2011: Workshop on User Modelling for Web Applications, Hong Kong, China.

Attuquayefio, S., & Addo, H. (2014). Review of studies with UTAUT as conceptual framework. *European Scientific Journal*, *10*(8).

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390‑412. https://doi.org/10.1016/j.jml.2007.12.005

Bakker, A. B. (2008a). The work-related flow inventory: Construction and initial validation of the WOLF. *Journal of Vocational Behavior*, *72*(3), 400‑414. https://doi.org/10.1016/j.jvb.2007.11.007

Bakker, A. B. (2008b). The work-related flow inventory: Construction and initial validation of the WOLF. *Journal of Vocational Behavior*, *72*(3), 400‑414. https://doi.org/10.1016/j.jvb.2007.11.007

Balog, A., & Pribeanu, C. (2010). The Role of Perceived Enjoyment in the Students' Acceptance of an Augmented Reality Teaching Platform: a Structural Equation Modelling Approach. *Studies in Informatics and Control*, *19*(3), 319‑330.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Studies*, *4*(3), 114–123.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574‑594. https://doi.org/10.1080/10447310802205776

Bargas-Avila, J. A., & Hornbaek, K. (2011). Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience. Dans *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2689–2698). New ork, N , USA : ACM. https://doi.org/10.1145/1978942.1979336

Bartneck, C., Duenser, A., Moltchanova, E., & Zawieska, K. (2015). Comparing the Similarity of Responses Received from Studies in Amazon's Mechanical Turk to Studies Conducted Online and with Direct Recruitment. *PLOS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0121595

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beaujean, A. A. (2014). *Latent Variable Modeling Using R: A Step-by-Step Guide*. New ork : Routledge.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, Decision Making and the Orbitofrontal Cortex. *Cerebral Cortex*, *10*(3), 295‑307. https://doi.org/10.1093/cercor/10.3.295

Benton, J. E., & Daly, J. L. (1991). A Question Order Effect in a Local Government Survey. *Public Opinion Quarterly*, *55*(4), 640‑642.

Berkman, M. I., & Karahoca, D. (2016). Re-assessing the Usability Metric for User Experience (UMUX) Scale. *J. Usability Studies*, *11*(3), 89–109.

Bhattacherjee, A., & Premkumar, G. (2004). Understanding Changes in Belief and Attitude toward Information Technology Usage: A Theoretical Model and Longitudinal Test. *MIS Quarterly*, *28*(2), 229‑254.

Bobillier-Chaumon, M., & Dubois, M. (2010). L'adoption des technologies en situation professionnelle : quelles articulations possibles entre acceptabilité et acceptation ? *Le travail humain*, *72*(4), 355‑382. https://doi.org/10.3917/th.724.0355

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual review of psychology*, *53*(1), 605–634.

Bone, P. F., Sharma, S., & Shimp, T. A. (1989). A Bootstrap Procedure for Evaluating Goodness-of-Fit Indices of Structural Equation and Confirmatory Factor Models. *Journal of Marketing Research*, *26*(1), 105. https://doi.org/10.2307/3172673

Bonnardel, N., Piolat, A., & Le Bigot, L. (2011). The impact of colour on Website appeal and users' cognitive processes. *Displays*, *32*, 69–80. https://doi.org/10.1016/j.displa.2010.12.002

Bosley, J. J. (2013). Creating a Short Usability Metric for User Experience (UMUX) Scale. *Interacting with Computers*, *25*(4), 317‑319. https://doi.org/10.1093/iwc/iwt007

Boyle, G. J. (1984). Reliability and validity of Izard's differential emotions scale. *Personality and Individual Differences*, *5*(6), 747‑750. https://doi.org/10.1016/0191-8869(84)90124-7

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49–59.

Brangier, É., Dufresne, A., & Hammes-Adelé, S. (2009). Approche symbiotique de la relation humain-technologie : perspectives pour l'ergonomie informatique. *Le travail humain*, *72*(4), 333‑353. https://doi.org/10.3917/th.724.0333

Brangier, É., Hammes-Adelé, S., & Bastien, J.-M. C. (2010). Analyse critique des approches de l'acceptation des technologies : de l'utilisabilité à la symbiose humain-technologie-organisation. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *60*(2), 129‑146. https://doi.org/10.1016/j.erap.2009.11.002

Brooke, J. (1996). SUS: A quick and dirty usability scale. Dans P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Éds), *Usability Evaluation In Industry* (pp. 189‑194). London : Taylor & Francis.

Brown, S. A., & Venkatesh, V. (2005). Model of Adoption of Technology in Households: A Baseline Model Test and Extension Incorporating Household Life Cycle. *MIS Quarterly*, *29*(3), 399‑426.

Brown, S. A., Venkatesh, V., & Goyal, S. (2012). Expectation Confirmation in Technology Use. *Information Systems Research*, *23*(2), 474‑487. https://doi.org/10.1287/isre.1110.0357

Bruner, G. C., & Kumar, A. (2005). Explaining consumer acceptance of handheld Internet devices. *Journal of Business Research*, *58*(5), 553‑558. https://doi.org/10.1016/j.jbusres.2003.08.002

Cairns, P. (2013). A Commentary on Short Questionnaires for Assessing Usability. *Interacting with Computers*, *25*(4), 312‑316. https://doi.org/10.1093/iwc/iwt019

Calisir, F., & Calisir, F. (2004). The relation of interface usability characteristics, perceived usefulness, and perceived ease of use to end-user satisfaction with enterprise resource planning (ERP) systems. *Computers in Human Behavior*, *20*(4), 505‑515. https://doi.org/10.1016/j.chb.2003.10.004

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81‑105.

Chapman, P., Selvarajah, S., & Webster, J. (1999). Engagement in multimedia training systems. Dans *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, 1999. HICSS-32*. https://doi.org/10.1109/HICSS.1999.772808

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's Corner: Testing Measurement Invariance of Second-Order Factor Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 471‑492. https://doi.org/10.1207/s15328007sem1203_7

Chen, L., Meservy, T., & Gillenson, M. (2012). Understanding Information Systems Continuance for Information-Oriented Mobile Applications. *Communications of the Association for Information Systems*, *30*(1).

Childers, T. L., Carr, C. L., Peck, J., & Carson, S. (2001). Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of Retailing*, *77*(4), 511‑535. https://doi.org/10.1016/S0022-4359(01)00056-2

Christophersen, T., & Konradt, U. (2011). Reliability, validity, and sensitivity of a single-item measure of online store usability. *International Journal of Human-Computer Studies*, *69*(4), 269‑280. https://doi.org/10.1016/j.ijhcs.2010.10.005

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300‑315. https://doi.org/10.1037/a0033805

Compeau, D. R., & Higgins, C. A. (1995). Computer Self-Efficacy: Development of a Measure and Initial Test. *MIS Quarterly*, *19*(2), 189‑211. https://doi.org/10.2307/249688

Corbitt, B. J., Thanasankit, T., & i, H. (2003). Trust and e-commerce: a study of consumer perceptions. *Electronic Commerce Research and Applications*, *2*(3), 203‑215. https://doi.org/10.1016/S1567-4223(03)00024-3

Cronin Jr., J. J., Brady, M. K., & Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing*, *76*(2), 193‑218. https://doi.org/10.1016/S0022-4359(00)00028-2

Csikszentmihalyi, M. (2014). The Concept of Flow. Dans M. Csikszentmihalyi, *Flow and the Foundations of Positive Psychology* (pp. 239‑263). (S.l.) : Springer Netherlands.

Cyr, D., Head, M., & Ivanov, A. (2006). Design aesthetics leading to m-loyalty in mobile commerce. *Information & Management*, *43*(8), 950‑963. https://doi.org/10.1016/j.im.2006.08.009

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors Predicting the Use of Technology: Findings From the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and aging*, *21*(2), 333‑352. https://doi.org/10.1037/0882-7974.21.2.333

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319‑340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982‑1003. https://doi.org/10.1287/mnsc.35.8.982

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology*, *22*(14), 1111‑1132. https://doi.org/10.1111/j.1559-1816.1992.tb00945.x

Davis, S., & Wiedenbeck, S. (2001). The mediating effects of intrinsic motivation, ease of use and usefulness perceptions on performance in first-time and subsequent computer users. *Interacting with Computers*, *13*(5), 549‑580. https://doi.org/10.1016/S0953-5438(01)00034-0

De Guinea, A. O., & Markus, M. L. (2009). Why Break the Habit of a Lifetime? Rethinking the Roles of Intention, Habit, and Emotion in Continuing Information Technology Use. *MIS Q.*, *33*(3), 433–444.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. (S.l.) : Springer Science & Business Media.

Desmet, P. (2002). *Designing emotions*. (S.l.) : (s.n.).

Desmet, P., Overbeeke, K., & Tax, S. (2001). Designing Products with Added Emotional Value: Development and Appllcation of an Approach for Research through Design. *The Design Journal*, *4*(1), 32‑47. https://doi.org/10.2752/146069201789378496

DeVellis, R. F. (2017). *Scale development: theory and applications* (Fourth edition). Los Angeles : SAGE.

Dewey, J. (2005). *Art as Experience*. New ork, N : TarcherPerigee.

Diefenbach, S., Kolb, N., & Hassenzahl, M. (2014). The « hedonic » in human-computer interaction: history, contributions, and future research directions (pp. 305‑314). ACM Press. https://doi.org/10.1145/2598510.2598549

Dillon, A., & Morris, M. G. (1996). User Acceptance of Information Technology: Theories and Models. *Annual Review of Information Science and Technology (ARIST)*, *31*, 3‑32.

Ding, X. D., Huang, ., & Verma, R. (2011). Customer experience in online financial services: A study of behavioral intentions for techno-ready market segments. *Journal of Service Management*, *22*(3), 344‑366. https://doi.org/10.1108/09564231111136863

Djamasbi, S., & Strong, D. M. (2008). The Effect of Positive Mood on Intention to Use Computerized Decision Aids. *Inf. Manage.*, *45*(1), 43–51. https://doi.org/10.1016/j.im.2007.10.002

Djamasbi, S., Strong, D. M., & Dishaw, M. (2010). Affect and acceptance: Examining the effects of positive mood on the technology acceptance model. *Decision Support Systems*, *48*(2), 383‑394. https://doi.org/10.1016/j.dss.2009.10.002

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. *Journal of Marketing Research*, *28*(3), 307. https://doi.org/10.2307/3172866

Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt, M. A. (2005). Reliability and validity of a single-item measure of job satisfaction. *American journal of health promotion: AJHP*, *19*(3), 194‑198.

Dubois, M., & Bobillier-Chaumon, M.-É. (2010). L'acceptabilité des technologies : bilans et nouvelles perspectives. *Le travail humain*, *72*(4), 305‑310. https://doi.org/10.3917/th.724.0305

Dwivedi, . K., Rana, N. P., Chen, H., & Williams, M. D. (2011). A Meta-analysis of the Unified Theory of Acceptance and Use of Technology (UTAUT). Dans M. Nüttgens, A. Gadatsch, K. Kautz, I. Schirmer, & N. Blinn (Éds), *Governance and Sustainability in Information Systems. Managing the Transfer and Diffusion of IT* (pp. 155‑170). (S.l.) : Springer Berlin Heidelberg.

Eisenberger, R., Jones, J. R., Stinglhamber, F., Shanock, L., & Randall, A. T. (2005). Flow experiences at work: for high need achievers alone? *Journal of Organizational Behavior*, *26*(7), 755‑775. https://doi.org/10.1002/job.337

Enders, C. K., & Bandalos, D. L. (2001). The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 430‑457. https://doi.org/10.1207/S15328007SEM0803_5

Erasmus, E., Rothmann, S., & van Eeden, C. (2015). A structural model of technology acceptance. *SA Journal of Industrial Psychology*, *41*(1), 12 pages. https://doi.org/10.4102/sajip.v41i1.1222

Éthier, J., Hadaya, P., Talbot, J., & Cadieux, J. (2008). Interface design and emotions experienced on B2C Web sites: Empirical testing of a research model. *Computers in Human Behavior*, *24*(6), 2771‑2791. https://doi.org/10.1016/j.chb.2008.04.004

Fan, X. (2003). Using Commonly Available Software For Bootstrapping In Both Substantive And Measurement Analyses. *Educational and Psychological Measurement*, *63*(1), 24‑50. https://doi.org/10.1177/0013164402239315

Festinger, L. (1962). *A Theory of Cognitive Dissonance*. (S.l.) : Stanford University Press.

Février, F. (2011). *Vers un modèle intégrateur « expérience-acceptation » : rôle des affects et de caractéristiques personnelles et contextuelles dans la détermination des intentions d'usage d'un environnement numérique de travail*. Rennes 2.

Février, F., Gauducheau, N., Jamet, E., Rouxel, G., & Salembier, P. (2011). La prise en compte des affects dans le domaine des interactions homme-machine : quels modèles, quelles méthodes, quels bénéfices ? *Le travail humain*, *74*(2), 183‑201. https://doi.org/10.3917/th.742.0183

Février, F., Jamet, E., & Rouxel, G. (2012). Vers une approche intégrative de l'adoption des technologies. Dans G. Dang Nguyen & P. Créach (Éds), *Actes du 9ème séminaire Marsouin* (pp. 219‑239). (S.l.) : (s.n.).

Février, F., Rouxel, G., & Jamet, E. (2012). Différences individuelles dans la formation des intentions d'usage d'un Environnement Numérique de Travail : déterminants personnologiques. Dans *Actes des XXèmes Journées Internationales de Psychologie Différentielle*. Université Rennes II.

Fiedler, M., Hossfeld, T., & Tran-Gia, P. (2010). A generic quantitative relationship between quality of experience and quality of service. *IEEE Network*, *24*(2), 36‑41. https://doi.org/10.1109/MNET.2010.5430142

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London ; Thousand Oaks, Calif : Sage.

Finstad, K. (2010). The Usability Metric for User Experience. *Interacting with Computers*, *22*(5), 323‑327. https://doi.org/10.1016/j.intcom.2010.04.004

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, Mass : Addison-Wesley Pub.

Fisher, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, *20*(2), 303. https://doi.org/10.1086/209351

Flavián, C., Guinalíu, M., & Gurrea, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information & Management*, *43*(1), 1‑14. https://doi.org/10.1016/j.im.2005.01.002

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286‑299. https://doi.org/10.1037/1040-3590.7.3.286

Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How Do Users Evaluate the Credibility of Web Sites?: A Study with over 2,500 Participants. Dans *Proceedings of the 2003 Conference on Designing for User Experiences* (pp. 1–15). New ork, N , USA : ACM. https://doi.org/10.1145/997078.997097

Fogg, B. J., & Tseng, H. (1999). The Elements of Computer Credibility. Dans *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 80–87). New ork, N , USA : ACM. https://doi.org/10.1145/302979.303001

Fu, B., Lin, J., Li, L., Faloutsos, C., Hong, J., & Sadeh, N. (2013). Why people hate your app: making sense of user feedback in a mobile app store (p. 1276). ACM Press. https://doi.org/10.1145/2487575.2488202

Fuchs, C., & Diamantopoulos, A. (2009). Using single-item measures for construct measurement in management research. *Die Betriebswirtschaft*, *69*(2), 195.

Gauducheau, N. (2009). Mesurer Les Emotions De L'Utilisateur: Quels Fondements Pour Une Démarche D'évaluation Des Systèmes Interactifs? Dans *Proceedings of the 21st International Conference on Association Francophone D'Interaction Homme-Machine* (pp. 183–192). Grenoble, France : ACM. https://doi.org/10.1145/1629826.1629855

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, *27*(1), 51‑90.

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). *Structural Equation Modeling And Regression: Guidelines For Research Practice*. (S.l.) : (s.n.).

Geng, X., Chen, Z., Lam, W., & Zheng, Q. (2013). Hedonic Evaluation over Short and Long Retention Intervals: The Mechanism of the Peak-End Rule: Hedonic Evaluation over Retention Intervals. *Journal of Behavioral Decision Making*, *26*(3), 225‑236. https://doi.org/10.1002/bdm.1755

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *The American Psychologist*, *59*(2), 93‑104. https://doi.org/10.1037/0003-066X.59.2.93

Gronier, G. (2016). Measuring the First Impression: Testing the Validity of the 5 Second Test. *J. Usability Studies*, *12*(1), 8–25.

Groth-Marnat, G. (2003). *Handbook of psychological assessment*. Hoboken, N.J. : John Wiley & Sons.

Gueorguieva, R., & Krystal, J. H. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry*, *61*(3), 310‑317. https://doi.org/10.1001/archpsyc.61.3.310

Ha, I., oon, ., & Choi, M. (2007). Determinants of adoption of mobile games under mobile broadband wireless access environment. *Information & Management*, *44*(3), 276‑286. https://doi.org/10.1016/j.im.2007.01.001

Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research*, *62*(5), 565‑571. https://doi.org/10.1016/j.jbusres.2008.06.016

Hagger, M. S. (2006). From Psychological Need Satisfaction to Intentional Behavior: Testing a Motivational Sequence in Two Behavioral Contexts. *Personality and Social Psychology Bulletin*, *32*(2), 131‑148. https://doi.org/10.1177/0146167205279905

Hall, R. H., & Hanna, P. (2004). The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention. *Behaviour & information technology*, *23*(3), 183–195.

Hall-Geisler, K. (2016). Waze and Esri make app-to-infrastructure possible. *TechCrunch*.

Hamm, J. A., & Hoffman, L. (2016). Working with Covariance: Using Higher-Order Factors in Structural Equation Modeling with Trust Constructs. Dans E. Shockley, T. M. S. Neal, L. M. PytlikZillig, & B. H. Bornstein (Éds), *Interdisciplinary Perspectives on Trust* (pp. 85‑97). (S.l.) : Springer International Publishing. https://doi.org/10.1007/978-3-319-22261-5_5

Harbich, S., & Hassenzahl, M. (2016). User Experience in the Work Domain: A Longitudinal Field Study. *Interacting with Computers*. https://doi.org/10.1093/iwc/iww022

Harris, L. R., & Brown, G. T. L. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data.

Hartmann, J., Sutcliffe, A., & Angeli, A. D. (2008). Towards a Theory of User Judgment of Aesthetics and User Interface Quality. *ACM Trans. Comput.-Hum. Interact.*, *15*(4), 15:1–15:30. https://doi.org/10.1145/1460355.1460357

Hassenzahl, M. (2004). The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction*, *19*(4), 319‑349. https://doi.org/10.1207/s15327051hci1904_2

Hassenzahl, M. (2005a). The Thing and I: Understanding the Relationship Between User and Product. Dans M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Éds), *Funology* (Vol. 3, pp. 31‑42). Dordrecht : Kluwer Academic Publishers.

Hassenzahl, M. (2005b). The Thing and I: Understanding the Relationship Between User and Product. Dans M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Éds), *Funology* (pp. 31‑42). (S.l.) : Springer Netherlands.

Hassenzahl, M. (2008). User Experience (UX): Towards an Experiential Perspective on Product Quality. Dans *Proceedings of the 20th International Conference of the Association Francophone D'Interaction Homme-Machine* (pp. 11–15). New ork, N , USA : ACM. https://doi.org/10.1145/1512714.1512717

Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, and interactive products–Facets of user experience. *Interacting with computers*, *22*(5), 353–362.

Hassenzahl, M., Kekez, R., & Burmester, M. (2002). The importance of a software's pragmatic quality depends on usage modes. Dans *Proceedings of the 6th international conference on Work With Display Units (WWDU 2002)* (pp. 275–276).

Hassenzahl, M., & Monk, A. (2010). The Inference of Perceived Usability From Beauty. *Human–Computer Interaction*, *25*(3), 235‑260. https://doi.org/10.1080/07370024.2010.500139

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, *25*(2), 91‑97. https://doi.org/10.1080/01449290500330331

Hayduk, L. A., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology*, *12*(1), 1‑17. https://doi.org/10.1186/1471-2288-12-159

He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on nowledge and Data Engineering*, *21*(9), 1263‑1284. https://doi.org/10.1109/TKDE.2008.239

Hole, L., & Williams, O. (2007). Gaining insight into the User eXperience (pp. 15‑18). Communication présentée au Towards a UX Manifesto, Law, Vermeeren, Hassenzahl & Blythe.

Hornbæk, K., & Law, E. L.-C. (2007). Meta-analysis of Correlations Among Usability Measures. Dans *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 617–626). New ork, N , USA : ACM. https://doi.org/10.1145/1240624.1240722

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). (S.l.) : John Wiley & Sons.

Hox, J., & Stoel, R. D. (2005). Multilevel and SEM Approaches to Growth Curve Modeling. Dans B. Everitt & D. C. Howell (Éds), *Encyclopedia of Statistics in Behavioral Science* (Vol. 3, pp. 1296‑1305). Chichester : John Wiley & Sons Ltd.

Hoyle, R. H. (2012). *Handbook of Structural Equation Modeling* (1re éd.). New ork : Guilford Press.

Hsu, C.-L., & Lin, J. C.-C. (2015). What drives purchase intention for paid mobile apps? – An expectation confirmation model with perceived value. *Electronic Commerce Research and Applications*, *14*(1), 46‑57. https://doi.org/10.1016/j.elerap.2014.11.003

Hsu, C.-L., & Lu, H.-P. (2004). Why Do People Play On-line Games? An Extended TAM with Social Influences and Flow Experience. *Inf. Manage.*, *41*(7), 853–868. https://doi.org/10.1016/j.im.2003.08.014

Hsu, C.-L., Lu, H.-P., & Hsu, H.-H. (2007). Adoption of the mobile Internet: An empirical study of multimedia message service (MMS). *Omega*, *35*(6), 715‑726. https://doi.org/10.1016/j.omega.2006.03.005

Hu, L., Bentler, P. M., & Kano, . (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, *112*(2), 351‑362. https://doi.org/10.1037/0033-2909.112.2.351

ISO. (2010). *ISO 9241-210:2010 - Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems* (1re éd.). (S.l.) : (s.n.).

ITU-T SG 12. (2007). Definition of quality of experience - COM12-LS 62-E, TD 109rev2 (PLEN/12). Geneva, Switzerland.

Jackson, J. D., i, M. ., & Park, J. S. (2013). An empirical test of three mediation models for the relationship between personal innovativeness and user acceptance of technology. *Information & Management*, *50*(4), 154‑161. https://doi.org/10.1016/j.im.2013.02.006

Jackson, S. A., & Eklund, R. C. (2002). Assessing Flow in Physical Activity: The Flow State Scale–2 and Dispositional Flow Scale–2. *Journal of Sport and Exercise Psychology*, *24*(2), 133‑150. https://doi.org/10.1123/jsep.24.2.133

Jackson, S. A., & Marsh, H. W. (1996). Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. *Journal of Sport and Exercise Psychology*, *18*(1), 17‑35. https://doi.org/10.1123/jsep.18.1.17

Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer trust in an Internet store. *Information Technology and Management*, *1*(1‑2), 45‑71. https://doi.org/10.1023/A:1019104520776

Jokinen, J. P. P. (2015). Emotional user experience: Traits, events, and states. *International Journal of Human-Computer Studies*, *76*, 67‑77. https://doi.org/10.1016/j.ijhcs.2014.12.006

Juhel, J. (2015). Modèles structuraux et inférence causale. Dans J. Juhel & G. Rouxel (Éds), *Différences et variabilités en psychologie* (pp. 197‑216). Rennes : P.U.R.

Kaiser, S., Wehrle, T., & Schenkel, K. (2013). Expression faciale. Dans D. Sander & K. R. Scherer, *Traité de psychologie des émotions* (Dunod). (S.l.) : (s.n.).

Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information Technology Adoption Across Time: A Cross-sectional Comparison of Pre-adoption and Post-adoption Beliefs. *MIS Q.*, *23*(2), 183–213. https://doi.org/10.2307/249751

Karapanos, E., Jain, J., & Hassenzahl, M. (2012). Theories, Methods and Case Studies of Longitudinal HCI Research. Dans *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (pp. 2727–2730). New ork, N , USA : ACM. https://doi.org/10.1145/2212776.2212706

Karapanos, E., Martens, J.-B., & Hassenzahl, M. (2010). On the Retrospective Assessment of Users' Experiences over Time: Memory or Actuality? Dans *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (pp. 4075–4080). New ork, N , USA : ACM. https://doi.org/10.1145/1753846.1754105

Karapanos, E., Martens, J.-B., & Hassenzahl, M. (2012). Reconstructing experiences with iScale. *International Journal of Human-Computer Studies*, *70*(11), 849‑865. https://doi.org/10.1016/j.ijhcs.2012.06.004

Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J.-B. (2009). User Experience over Time: An Initial Framework. Dans *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 729–738). New ork, N , USA : ACM. https://doi.org/10.1145/1518701.1518814

Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J.-B. (2010). Measuring the dynamics of remembered experience over time. *Interacting with Computers*, *22*(5), 328‑335. https://doi.org/10.1016/j.intcom.2010.04.003

Kay, R. H., & Loverock, S. (2008). Assessing emotions related to learning new software: The computer emotion scale. *Computers in Human Behavior*, *24*(4), 1605‑1623. https://doi.org/10.1016/j.chb.2007.06.002

Keltner, D., & Gross, J. J. (1999). Functional Accounts of Emotions. *Cognition & Emotion*, *13*(5), 467‑480. https://doi.org/10.1080/026999399379140

Kenny, D. A., & Milan, S. (2012). Identification: A non-technical discussion of a technical issue. Dans R. H. Hoyle (Éd.), *Handbook of Structural Equation Modeling* (1re éd., pp. 145‑163). New ork : Guilford Press.

Khechine, H., Lakhal, S., & Ndjambou, P. (2016). A meta-analysis of the UTAUT model: Eleven years later: A meta-analysis of the UTAUT model: Eleven years later. *Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration*, *33*(2), 138‑152. https://doi.org/10.1002/cjas.1381

Kim, J. (2015). The Three Threads of Experience. Dans *Design for Experience* (pp. 23‑40). (S.l.) : Springer, Cham. https://doi.org/10.1007/978-3-319-14304-0_2

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, *43*(6), 740‑755. https://doi.org/10.1016/j.im.2006.05.003

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263‑291. https://doi.org/10.1007/BF00993889

Klem, A. M., & Connell, J. P. (2004). Relationships Matter: Linking Teacher Support to Student Engagement and Achievement. *Journal of School Health*, *74*(7), 262‑273. https://doi.org/10.1111/j.1746-1561.2004.tb08283.x

Kline, R. B. (2010). *Principles and Practice of Structural Equation Modeling, Third Edition* (3rd edition). New ork : The Guilford Press.

Koufaris, M. (2002). Applying the Technology Acceptance Model and Flow Theory to Online Consumer Behavior. *Information Systems Research*, *13*(2), 205‑223. https://doi.org/10.1287/isre.13.2.205.83

Krantz, J. H., & Dalal, R. (2000). Validity of Web-Based Psychological Research. Dans *Psychological Experiments on the Internet* (pp. 35‑60). (S.l.) : Elsevier. https://doi.org/10.1016/B978-012099980-4/50003-4

Kuhr, C. (2013). *Measuring the User Experience of Mobile Applications – an Empirical Validation of a Quantitative Method*. Berlin. Repéré à http://mecue.de/Homepage%20Content/02%20Links%20%26%20Literatur/MA_Kuhr.pdf

Kujala, S., & Nurkka, P. (2012). Sentence completion for evaluating symbolic meaning. *International Journal of Design*, *6*(3).

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., & Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. *Interacting with Computers*, *23*(5), 473‑483. https://doi.org/10.1016/j.intcom.2011.06.005

Kuniavsky, M. (2009). User Experience and HCI. Dans A. Sears & J. A. Jacko (Éds), *Human-Computer Interaction* (CRC Press, pp. 3‑22). Boca Raton, FL, USA : (s.n.).

Kuutti, K. (2010). Where Are the Ionians of User Experience Research? Dans *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 715–718). New ork, N , USA : ACM. https://doi.org/10.1145/1868914.1869012

Kwon, K. H., & Chon, B. S. (2009). Social Influences on Terrestrial and Satellite Mobile-TV Adoption in Korea: Affiliation, Positive Self-Image, and Perceived Popularity. *International Journal on Media Management*, *11*(2), 49‑60. https://doi.org/10.1080/14241270902756419

Lallemand, C., & Gronier, G. (2015). *Méthodes de design UX. 30 méthodes fondamentales pour concevoir et évaluer les systèmes interactifs*. (S.l.) : E ROLLES.

Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, *43*, 35‑48. https://doi.org/10.1016/j.chb.2014.10.048

Lallemand, C., Koenig, V., Gronier, G., & Martin, R. (2015). Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *65*(5), 239‑252. https://doi.org/10.1016/j.erap.2015.08.002

Larsen, K. R., Nevo, D., & Rich, E. (2008). Exploring the Semantic Validity of Questionnaire Scales. Dans *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual* (pp. 440‑440). https://doi.org/10.1109/HICSS.2008.165

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. Dans *HCI and Usability for Education and Work* (pp. 63‑76). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-89350-9_6

Law, E. L.-C., & van Schaik, P. (2010). Modelling user experience – An agenda for research and practice. *Interacting with Computers*, *22*(5), 313‑322. https://doi.org/10.1016/j.intcom.2010.04.006

Lee, M. K. O., Cheung, C. M. K., & Chen, Z. (2005). Acceptance of Internet-based learning medium: the role of extrinsic and intrinsic motivation. *Information & Management*, *42*(8), 1095‑1104. https://doi.org/10.1016/j.im.2003.10.007

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management*, *40*(3), 191‑204. https://doi.org/10.1016/S0378-7206(01)00143-4

Lehmann, J., Lalmas, M., om-Tov, E., & Dupret, G. (2012). Models of User Engagement. Dans *Proceedings of the 20th International Conference on User Modeling, Adaptation, and Personalization* (pp. 164–175). Berlin, Heidelberg : Springer-Verlag. https://doi.org/10.1007/978-3-642-31454-4_14

Lei, P.-W., & Wu, Q. (2007). Introduction to Structural Equation Modeling: Issues and Practical Considerations. *Educational Measurement: Issues and Practice*, *26*(3), 33‑43. https://doi.org/10.1111/j.1745-3992.2007.00099.x

Lerner, J. S., Li, ., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, *66*(1), 799‑823. https://doi.org/10.1146/annurev-psych-010213-115043

Lewis, J. R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *Int. J. Hum.-Comput. Interact.*, *7*(1), 57–78. https://doi.org/10.1080/10447319509526110

Lewis, J. R. (2012). Usability Testing. Dans G. Salvendy (Éd.), *Handbook of Human Factors and Ergonomics* (pp. 1267‑1312). (S.l.) : John Wiley & Sons, Inc.

Lewis, J. R. (2013). Critical Review of « The Usability Metric for User Experience ». *Interacting with Computers*, *25*(4), 320‑324. https://doi.org/10.1093/iwc/iwt013

Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (Éds). (2010). *Handbook of Emotions, Third Edition* (3rd edition). New ork : The Guilford Press.

Lim, ., Donaldson, J., Jung, H., Kunz, B., Royer, D., Ramalingam, S., … Stolterman, E. (2008). Emotional Experience and Interaction Design. Dans C. Peter & R. Beale (Éds), *Affect and Emotion in Human-Computer Interaction* (pp. 116‑129). (S.l.) : Springer Berlin Heidelberg.

Limayem, M., Hirt, S. G., & Cheung, C. M. K. (2007). How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance. *MIS Quarterly*, *31*(4), 705‑737.

Lindgaard, G., & Dudek, C. (2003). What is this evasive beast we call user satisfaction? *Interacting with Computers*, *15*(3), 429‑452. https://doi.org/10.1016/S0953-5438(02)00063-2

Luminet, O. (2008). *Psychologie des émotions: Confrontation et évitement*. (S.l.) : De Boeck Supérieur.

Luo, X. (2002). Uses and Gratifications Theory and E-Consumer Behaviors: A Structural Equation Modeling Study. *Journal of Interactive Advertising*, *2*(2), 34‑41. https://doi.org/10.1080/15252019.2002.10722060

Lynch, S. M. (Éd.). (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New ork, N : Springer New ork. https://doi.org/10.1007/978-0-387-71265-9

Ma, Q., & Liu, L. (2004). The Technology Acceptance Model: A Meta-Analysis of Empirical Findings. *Journal of Organizational and End User Computing*, *16*(1), 59‑72. https://doi.org/10.4018/joeuc.2004010104

Mahlke, S. (2007). Aesthetic and Symbolic Qualities as Antecedents of Overall Judgements of Interactive Products. Dans N. B.-K. Bs. AKC MSc,, PGCAP, A. B. B. CITP MA,, MBCS, CEng, P. C. B. PGCertHE MA, & L. N. Bs. MSc (Éds), *People and Computers XX — Engage* (pp. 57‑64). (S.l.) : Springer London.

Mahlke, S., Lemke, I., & Thüring, M. (2007). The Diversity of Non-instrumental Qualities in Human-Technology Interaction. *MMI Interaktiv*, *13*, 55‑64.

Mahlke, S., & Lindgaard, G. (2007). Emotional Experiences and Quality Perceptions of Interactive Products. Dans J. A. Jacko (Éd.), *Human-Computer Interaction. Interaction Design and Usability* (pp. 164‑173). (S.l.) : Springer Berlin Heidelberg.

Mahlke, S., & Minge, M. (2008). Consideration of Multiple Components of Emotions in Human-Technology Interaction. Dans C. Peter & R. Beale (Éds), *Affect and Emotion in Human-Computer Interaction* (pp. 51‑62). (S.l.) : Springer Berlin Heidelberg.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519‑530. https://doi.org/10.1093/biomet/57.3.519

Maricutoiu, L. P. (2014). A Meta-analysis on the Antecedents and Consequences of Computer Anxiety. *Procedia - Social and Behavioral Sciences*, *127*, 311‑315. https://doi.org/10.1016/j.sbspro.2014.03.262

Martin, N., Erhel, S., Jamet, É., & Rouxel, G. (2015). What Links Between User Experience and Acceptability? Dans *Proceedings of the 27th Conference on L'Interaction Homme-Machine* (p. 34:1–34:6). New ork, N , USA : ACM. https://doi.org/10.1145/2820619.2825015

Martin, N., Jamet, É., Erhel, S., & Rouxel, G. (2016a). Evaluation of technological products in mobility context: first steps toward new scales with single items (pp. 237‑245). ACM Press. https://doi.org/10.1145/3004107.3004133

Martin, N., Jamet, É., Erhel, S., & Rouxel, G. (2016b). From Acceptability to Acceptance: Does Experience with the Product Influence User Initial Representations? Dans C. Stephanidis (Éd.), *HCI International 2016 – Posters' Extended Abstracts* (Vol. 617, pp. 128‑133). Cham : Springer International Publishing.

McCarthy, J. J., & Wright, P. (2007). *Technology as Experience*. Cambridge, Mass. : The MIT Press.

McCrae, R. R., Kurtz, J. E., amagata, S., & Terracciano, A. (2011). Internal Consistency, Retest Reliability, and Their Implications for Personality Scale Validity. *Personality and Social Psychology Review*, *15*(1), 28‑50. https://doi.org/10.1177/1088868310366253

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525‑543. https://doi.org/10.1007/BF02294825

Microsoft Corporation. (2007). *Quality of Experience - A Strategic Competitive Advantage of Microsoft Unified Communications*.

Milfont, T. L., & Fischer, R. (2015). Testing measurement invariance across groups: applications in cross-cultural research. *International Journal of Psychological Research*, *3*(1), 111‑130. https://doi.org/10.21500/20112084.857

Minge, M. (2013). meCUE. *meCUE*. Repéré à http://mecue.de/english/index.html

Minge, M., Thüring, M., & Wagner, I. (2016). Developing and Validating an English Version of the meCUE Questionnaire for Measuring User Experience. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 2063‑2067. https://doi.org/10.1177/1541931213601468

Mitra, K., Zaslavsky, A., & Åhlund, C. (2011). Dynamic Bayesian Networks for Sequential Quality of Experience Modelling and Measurement. Dans S. Balandin, . Koucheryavy, & H. Hu (Éds), *Smart Spaces and Next Generation Wired/Wireless Networking* (pp. 135‑146). (S.l.) : Springer Berlin Heidelberg. Repéré à http://link.springer.com/chapter/10.1007/978-3-642-22875-9_12

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal Theories of Emotion: State of the Art and Future Development. *Emotion Review*, *5*(2), 119‑124. https://doi.org/10.1177/1754073912468165

Morse, J. M. (2009). *Preventing patient falls: establishing a fall intervention program* (2nd ed). New ork : Springer Pub. Co.

Mortel, V. D., & Francisca. (2008). Faking it: Social desirability response bias in self-report research, *25*(4), 40‑48.

Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, *68*(10), 689‑709. https://doi.org/10.1016/j.ijhcs.2010.05.006

Muenchen, R. A. (2013). The Popularity of Data Science Software. *r4stats.com*. Repéré à http://r4stats.com/articles/popularity/

Nakamura, J., & Csikszentmihalyi, M. (2002). The concept of flow. Dans C. R. Snyder & S. J. Lopez (Éds), *Handbook of positive psychology* (pp. 89‑105). New ork, N , US : Oxford University Press.

Nakazawa, M. (2017). *fmsb: Functions for Medical Statistics Book with some Demographic Data*. (S.l.) : (s.n.). Repéré à https://CRAN.R-project.org/package=fmsb

NF EN ISO 9241-11. (1998). *Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) - Partie 11 : Ligne directrice concernant l'utilisabilité*. (S.l.) : (s.n.).

NF EN ISO 9241-210. (2010). *Ergonomie de l'interaction homme-système - Partie 210 : Conception centrée sur l'opérateur humain pour les systèmes interactifs*. (S.l.) : (s.n.).

Nielsen, J. (1994). *Usability Engineering* (New edition). San Francisco, Calif. : Morgan Kaufmann.

Oberski, D. L., & Satorra, A. (2013). Measurement Error Models With Uncertainty About the Error Variance. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(3), 409‑428. https://doi.org/10.1080/10705511.2013.797820

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938‑955. https://doi.org/10.1002/asi.20801

O'Brien, H. L., & Toms, E. G. (2010). The Development and Evaluation of a Survey to Measure User Engagement. *J. Am. Soc. Inf. Sci. Technol.*, *61*(1), 50–69. https://doi.org/10.1002/asi.v61:1

O'Brien, H. L., & Toms, E. G. (2013). Examining the generalizability of the User Engagement Scale (UES) in exploratory search. *Information Processing & Management*, *49*(5), 1092‑1107. https://doi.org/10.1016/j.ipm.2012.08.005

Obrist, M., Roto, V., Vermeeren, A., Väänänen-Vainio-Mattila, K., Law, E. L.-C., & Kuutti, K. (2012). In Search of Theoretical Foundations for UX Research and Practice. Dans *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (pp. 1979–1984). New ork, N , USA : ACM. https://doi.org/10.1145/2212776.2223739

OCDE. (2003). *Seizing the benefits of ICT in a digital economy*.

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, *17*(4), 460. https://doi.org/10.2307/3150499

Ortega Egea, J. M., & Román González, M. V. (2011). Explaining physicians' acceptance of EHCR systems: An extension of TAM with trust and risk factors. *Computers in Human Behavior*, *27*(1), 319‑332. https://doi.org/10.1016/j.chb.2010.08.010

Overby, J. W., & Lee, E.-J. (2006). The effects of utilitarian and hedonic online shopping value on consumer preference and intentions. *Journal of Business Research*, *59*(10–11), 1160‑1166. https://doi.org/10.1016/j.jbusres.2006.03.008

Pakanen, M. (2015). *Visual design examples in the evaluation of anticipated user experience at the early phases of research and development*. University of Oulu.

Pavlou, P. A. (2003). Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. *International Journal of Electronic Commerce*, *7*(3), 101‑134.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148‑158. https://doi.org/10.1038/nrn2317

Porat, T., & Tractinsky, N. (2012). It's a Pleasure Buying Here: The Effects of Web-Store Design on Consumers' Emotions and Attitudes. *Human–Computer Interaction*, *27*(3), 235‑276. https://doi.org/10.1080/07370024.2011.646927

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria : R Foundation for Statistical Computing.

Raita, E., & Oulasvirta, A. (2011). Too good to be bad: Favorable product expectations boost subjective usability ratings. *Interacting with Computers*, *23*(4), 363‑371. https://doi.org/10.1016/j.intcom.2011.04.002

Raykov, T., & Marcoulides, G. A. (1999). On desirability of parsimony in structural equation model selection. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(3), 292‑300. https://doi.org/10.1080/10705519909540135

Revelle, W. (2015). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois : Northwestern University. (R package version 1.5.4).

Riegelsberger, J., Sasse, M. A., & McCarthy, J. D. (2005). The mechanics of trust: A framework for research and design. *International Journal of Human-Computer Studies*, *62*(3), 381‑422. https://doi.org/10.1016/j.ijhcs.2005.01.001

Rindskopf, D., & Rose, T. (1988). Some Theory and Applications of Confirmatory Second-Order Factor Analysis. *Multivariate Behavioral Research*, *23*(1), 51‑67. https://doi.org/10.1207/s15327906mbr2301_3

Robert, J.-M., & Lesage, A. (2011a). Designing and evaluating user experience. Dans G. A. Boy (Éd.), *Handbook of Human-Computer Interaction* (pp. 321‑338). Ashgate, U.K. : (s.n.).

Robert, J.-M., & Lesage, A. (2011b). From usability to user experience with interactive systems. Dans G. A. Boy (Éd.), *Handbook of Human-Computer Interaction* (pp. 303–320). Ashgate, U.K. : (s.n.).

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *27*(2), 151‑161. https://doi.org/10.1177/0146167201272002

Rogers, E. M. (1995). *Diffusion of Innovations, Fourth Edition* (4 edition). New ork : Free Press.

Rönkkö, M., & Evermann, J. (2013). A Critical Examination of Common Beliefs About Partial Least Squares Path Modeling. *Organizational Research Methods*, *16*(3), 425‑448. https://doi.org/10.1177/1094428112474693

Rosseel, . (2012a). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Rosseel, . (2012b). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Roto, V. (2007). User experience from product creation perspective. Dans *Towards a UX Manifesto* (pp. 31‑34). Law, Vermeeren, Hassenzahl & Blythe.

Rozendaal, M. (M. C. ., Keyson, D. V., & Ridder, H. (2009). Product Features and Task Effects on Experienced Richness, Control and Engagement in Voicemail Browsing. *Personal Ubiquitous Comput.*, *13*(5), 343–354. https://doi.org/10.1007/s00779-008-0201-8

Ryan, G. S. (2013). Online social networks for patient involvement and recruitment in clinical research. *Nurse Researcher*, *21*(1), 35‑39. https://doi.org/10.7748/nr2013.09.21.1.35.e302

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54‑67. https://doi.org/10.1006/ceps.1999.1020

Saariluomaand, P., & Jokinen, J. P. P. (2014). Emotional Dimensions of User Experience: A User Psychological Analysis. *International Journal of Human-Computer Interaction*, *30*(4), 303‑320. https://doi.org/10.1080/10447318.2013.858460

Sander, D., & Scherer, K. R. (2014). La psychologie des émotions : survol des théories et débats essentiels. Dans *Traité de psychologie des émotions* (Dunod, pp. 1‑39). (S.l.) : (s.n.).

Schade, J., & Schlag, B. (2003). Acceptability of urban transport pricing strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *6*(1), 45‑61. https://doi.org/10.1016/S1369-8478(02)00046-3

Schaffer, S., Seebode, J., Wechsung, I., Metze, F., & Kühnel, C. (2009). Influence of User Characteristics on the Usage of Gesture and Speech in a Smart Office Environment. Dans *In Proceedings of the 8th International Gesture Workshop 2009*.

Schaufeli, W. B., Salanova, M., González-romá, V., & Bakker, A. B. (2002). The Measurement of Engagement and Burnout: A Two Sample Confirmatory Factor Analytic Approach. *Journal of Happiness Studies*, *3*(1), 71‑92. https://doi.org/10.1023/A:1015630930326

Schepers, J., & Wetzels, M. (2007). A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. *Information & Management*, *44*(1), 90‑103. https://doi.org/10.1016/j.im.2006.10.007

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695‑729. https://doi.org/10.1177/0539018405058216

Schuitema, G., Steg, L., & Forward, S. (2010). Explaining differences in acceptability before and acceptance after the implementation of a congestion charge in Stockholm. *Transportation Research Part A: Policy and Practice*, *44*(2), 99‑109. https://doi.org/10.1016/j.tra.2009.11.005

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology*, *21*(2), 277‑287. https://doi.org/10.1002/acp.1340

Scott, J. E., & Walczak, S. (2009). Cognitive engagement with a multimedia ERP training tool: Assessing computer self-efficacy and technology acceptance. *Information & Management*, *46*(4), 221‑232. https://doi.org/10.1016/j.im.2008.10.003

Seckler, M., Heinz, S., Forde, S., Tuch, A. N., & Opwis, K. (2015). Trust and distrust on the web: User experiences and website characteristics. *Computers in Human Behavior*, *45*, 39‑50. https://doi.org/10.1016/j.chb.2014.11.064

Sezgin, E., & ıldırım, S. Ö. (2016). Trends of Factors and Theories in Health Information Systems Acceptance: 2002 – 2014 Review. *http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-9978-6.ch085*, 1085‑1104. https://doi.org/10.4018/978-1-4666-9978-6.ch085

Shih, H.-P. (2004). An empirical study on predicting user acceptance of e-shopping on the Web. *Information & Management*, *41*(3), 351‑368. https://doi.org/10.1016/S0378-7206(03)00079-X

Song, F., Parekh, S., Hooper, L., Loke, . K., Ryder, J., Sutton, A. J., … Harvey, I. (2010). Dissemination and publication of research findings: an updated review of related biases. *Health Technology Assessment (Winchester, England)*, *14*(8), iii, ix‑xi, 1-193. https://doi.org/10.3310/hta14080

Sorgenfrei, C., Ebner, K., Smolnik, S., & Jennex, M. E. (2014). From Acceptance To Outcome: Towards An Integrative Framework For Information Technology Adoption. *ECIS 2014 Proceedings*.

Suhr, D. D. (2005). Principal component analysis vs. exploratory factor analysis. *SUGI 30 proceedings*, 203‑230.

Šumak, B., & Šorgo, A. (2016). The acceptance and use of interactive whiteboards among teachers: Differences in UTAUT determinants between pre- and post-adopters. *Computers in Human Behavior*, *64*, 602‑620. https://doi.org/10.1016/j.chb.2016.07.037

Sun, H., & Zhang, P. (2006). Causal Relationships between Perceived Enjoyment and Perceived Ease of Use: An Alternative Approach. *Journal of the Association for Information Systems*, *7*(9).

Sun, H., & Zhang, P. (2008). An exploration of affect factors and their role in user technology acceptance: Mediation and causality. *Journal of the American Society for Information Science and Technology*, *59*(8), 1252‑1263. https://doi.org/10.1002/asi.20819

Swann, C., Keegan, R. J., Piggott, D., & Crust, L. (2012). A systematic review of the experience, occurrence, and controllability of flow states in elite sport. *Psychology of Sport and Exercise*, *13*(6), 807‑819. https://doi.org/10.1016/j.psychsport.2012.05.006

Taiwo, . A., & Downe, A. G. (2013). The Theory of user Acceptance and Use of Technology (UTAUT): A meta-analytic review of empirical findings. *Journal of Theoretical & Applied Information Technology*, *49*(1).

Taylor, S., & Todd, P. (1995a). Assessing IT Usage: The Role of Prior Experience. *MIS Quarterly*, *19*(4), 561‑570. https://doi.org/10.2307/249633

Taylor, S., & Todd, P. A. (1995b). Understanding Information Technology Usage: A Test of Competing Models. *Information Systems Research*, *6*(2), 144‑176. https://doi.org/10.1287/isre.6.2.144

Tenenhaus, M., Vinzi, V. E., Chatelin, .-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis*, *48*(1), 159‑205. https://doi.org/10.1016/j.csda.2004.03.005

Teo, T. S. H., Lim, V. K. G., & Lai, R. . C. (1999). Intrinsic and extrinsic motivation in Internet usage. *Omega*, *27*(1), 25‑37. https://doi.org/10.1016/S0305-0483(98)00028-0

Thayer, R. E. (2010). *The biopsychology of mood and arousal*. New ork, N. : Oxford University Press.

Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal Computing: Toward a Conceptual Model of Utilization. *MIS Quarterly*, *15*(1), 125‑143. https://doi.org/10.2307/249443

Thüring, M., & Mahlke, S. (2007a). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, *42*(4), 253‑264. https://doi.org/10.1080/00207590701396674

Thüring, M., & Mahlke, S. (2007b). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, *42*(4), 253–264.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, *13*(2), 127‑145. https://doi.org/10.1016/S0953-5438(00)00031-X

Tractinsky, N., & Zmiri, D. (2008). Exploring Attributes of Skins as Potential Antecedents of Emotion in HCI. Dans P. A. Fishwick (Éd.), *Aesthetic Computing* (pp. 405‑422). (S.l.) : MIT Press.

Tricot, A., Plégat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G., & Morcillo, A. (2003). Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH (pp. 391‑402). Communication présentée au Environnements Informatiques pour l'Apprentissage Humain 2003, ATIEF ; INRP. Repéré à https://edutice.archives-ouvertes.fr/edutice-00000154/document

Tuch, A. N., Roth, S. P., Hornbæk, K., Opwis, K., & Bargas-Avila, J. A. (2012). Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behavior*, *28*(5), 1596‑1607. https://doi.org/10.1016/j.chb.2012.03.024

van der Heijden, H. (2003). Factors influencing the usage of websites: the case of a generic portal in The Netherlands. *Information & Management*, *40*(6), 541‑549. https://doi.org/10.1016/S0378-7206(02)00079-4

van Schaik, P., & Ling, J. (2008). Modelling user experience with web sites: Usability, hedonic value, beauty and goodness. *Interacting with Computers*, *20*(3), 419‑432. https://doi.org/10.1016/j.intcom.2008.03.001

van Schaik, P., & Ling, J. (2011). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. *Interacting with Computers*, *23*(1), 18‑32. https://doi.org/10.1016/j.intcom.2010.07.002

van Schaik, P., Zeng, ., & Spears, I. (2014). Using Dynamic Bayesian Networks to Model User-Experience. Dans L. Cao, . Zeng, A. L. Symeonidis, V. Gorodetsky, J. P. Müller, & P. S. u (Éds), *Agents and Data Mining Interaction* (pp. 3‑13). (S.l.) : Springer Berlin Heidelberg.

Venkatesh, V., & Davis, F. D. (1996). A Model of the Antecedents of Perceived Ease of Use: Development and Test. *Decision Sciences*, *27*(3), 451‑481. https://doi.org/10.1111/j.1540-5915.1996.tb00860.x

Venkatesh, V., Davis, F. D., & Morris, M. (2007). Dead Or Alive? The Development, Trajectory And Future Of Technology Adoption Research. *Journal of the Association for Information Systems*, *8*(4), 267‑286.

Venkatesh, V., Morris, M., Davis, G., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *Management Information Systems Quarterly*, *27*(3).

Venkatesh, V., & Morris, M. G. (2000). Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior. *MIS Quarterly*, *24*(1), 115‑139. https://doi.org/10.2307/3250981

Venkatesh, V., & Speier, C. (1999). Computer Technology Training in the Workplace: A Longitudinal Investigation of the Effect of Mood. *Organizational Behavior and Human Decision Processes*, *79*(1), 1‑28. https://doi.org/10.1006/obhd.1999.2837

Venkatesh, V., Speier, C., & Morris, M. G. (2002). User acceptance enablers in individual decision making about technology: Toward an integrated model. *Decision sciences*, *33*(2), 297–316.

Venkatesh, V., Thong, J. . L., Chan, F. K. ., Hu, P. J.-H., & Brown, S. A. (2011). Extending the two-stage information systems continuance model: incorporating UTAUT predictors and the role of context. *Information Systems Journal*, *21*(6), 527‑555. https://doi.org/10.1111/j.1365-2575.2011.00373.x

Venkatesh, V., Thong, J. . L., & Xu, X. (2012a). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Q.*, *36*(1), 157–178.

Venkatesh, V., Thong, J. . L., & Xu, X. (2012b). *Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology* (Rapport No. ID 2002388). Rochester, N : Social Science Research Network.

von Wilamowitz-Moellendorff, M., Hassenzahl, M., & Platz, A. (2006). Dynamics of user experience: How the perceived quality of mobile phones changes over time. Dans *User Experience - Towards a unified view, Workshop at the 4th Nordic Conference on Human-Computer Interaction* (pp. 74‑78). (S.l.) : (s.n.).

Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude. *Journal of Marketing Research*, *40*(3), 310‑320.

Wang, . J., Hong, S., & Lou, H. (2010). Beautiful beyond useful? The role of web aesthetics. *The Journal of Computer Information Systems*, *50*(3), 121.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063‑1070.

Webster, J., & Ho, H. (1997). Audience Engagement in Multimedia Presentations. *SIGMIS Database*, *28*(2), 63–77. https://doi.org/10.1145/264701.264706

Wechsung, I., & De Moor, K. (2014). Quality of Experience versus User Experience. Dans S. Möller & A. Raake (Éds), *Quality of Experience: Advanced Concepts, Applications and Method* (pp. 35‑53). (S.l.) : (s.n.).

Weiss, H. M., Nicholas, J. P., & Daus, C. S. (1999). An Examination of the Joint Effects of Affective Experiences and Job Beliefs on Job Satisfaction and Variations in Affective Experiences over Time. *Organizational Behavior and Human Decision Processes*, *78*(1), 1‑24. https://doi.org/10.1006/obhd.1999.2824

White, C., & u, . (2005). Satisfaction emotions and consumer behavioral intentions. *Journal of Services Marketing*, *19*(6), 411‑420. https://doi.org/10.1108/08876040510620184

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. (S.l.) : Springer-Verlag New ork. Repéré à http://ggplot2.org

Wimmer, B., Wöckl, B., Leitner, M., & Tscheligi, M. (2010). Measuring the Dynamics of User Experience in Short Interaction Sequences. Dans *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 825–828). New ork, N , USA : ACM. https://doi.org/10.1145/1868914.1869039

Wright, P., McCarthy, J. J., & Meekison, L. (2005). Making Sense of Experience. Dans M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Éds), *Funology* (pp. 43‑53). (S.l.) : Springer Netherlands.

Wright, P., Wallace, J., & McCarthy, J. (2008). Aesthetics and Experience-centered Design. *ACM Trans. Comput.-Hum. Interact.*, *15*(4), 18:1–18:21. https://doi.org/10.1145/1460355.1460360

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research and Evaluation*, *12*(3), 1–26.

Wu, W., Arefin, A., Rivas, R., Nahrstedt, K., Sheppard, R., & ang, Z. (2009). Quality of Experience in Distributed Interactive Multimedia Environments: Toward a Theoretical Framework. Dans *Proceedings of the 17th ACM International Conference on Multimedia* (pp. 481–490). New ork, N , USA : ACM. https://doi.org/10.1145/1631272.1631338

Wyatt, J. C. (2000). When to Use Web-based Surveys. *Journal of the American Medical Informatics Association : JAMIA*, *7*(4), 426‑429.

ang, S., Lu, ., Gupta, S., Cao, ., & Zhang, R. (2012). Mobile payment services adoption across time: An empirical study of the effects of behavioral beliefs, social influences, and personal traits. *Computers in Human Behavior*, *28*(1), 129‑142. https://doi.org/10.1016/j.chb.2011.08.019

ang, T., Linder, J., & Bolchini, D. (2012). DEEP: Design-Oriented Evaluation of Perceived Usability. *International Journal of Human-Computer Interaction*, *28*(5), 308‑346. https://doi.org/10.1080/10447318.2011.586320

eh, R. K.-J., & Teng, J. T. C. (2012). Extended conceptualisation of perceived usefulness: empirical test in the context of information system use continuance. *Behaviour & Information Technology*, *31*(5), 525‑540. https://doi.org/10.1080/0144929X.2010.517272

ogasara, T., Popovic, V., Kraal, B. J., & Chamorro-Koc, M. (2011). General characteristics of anticipated user experience (AUX) with interactive products. Dans N. Roozenburg, L.-L. Chen, & P. J. Stappers (Éds), *Proceedings of IASDR2011 : the 4th World Conference on Design Research : Diversity and Unity* (pp. 1‑11). Delft University of Technology, Delft, The Netherlands : IASDR.

oo, S. J., Han, S., & Huang, W. (2012). The roles of intrinsic motivators and extrinsic motivators in promoting e-learning in the workplace: A case from South Korea. *Computers in Human Behavior*, *28*(3), 942‑950. https://doi.org/10.1016/j.chb.2011.12.015

uan, P., Bare, M. G., Johnson, M. O., & Saberi, P. (2014). Using Online Social Media for Recruitment of Human Immunodeficiency Virus-Positive Participants: A Cross-Sectional Survey. *Journal of Medical Internet Research*, *16*(5). https://doi.org/10.2196/jmir.3229

Annexes

Annexe 1. Questionnaire TAM (F. D. Davis, 1989)

Perception de Facilité d'Usage

- Apprendre à utiliser le système serait facile pour moi
- Je trouve facile de faire faire au système ce que je veux qu'il fasse
- Il me serait facile de devenir compétent dans l'utilisation de ce système
- Je trouverais facile d'utiliser ce système
- L'utilisation de ce dispositif est entièrement sous mon contrôle
- Je serai capable d'utiliser ce système
- J'ai les compétences, les connaissances, et la capacité à utiliser ce système

Utilité perçue

- Utiliser ce système dans mon travail me permettrait d'accomplir mes tâches plus rapidement
- Utiliser ce système me permettrait d'augmenter mes performances au travail
- Utiliser ce système augmenterait ma productivité
- Utiliser ce système améliorerait mon efficacité au travail
- Utiliser ce système rendrait mon travail plus facile
- Je trouverai ce système utile dans mon travail

Annexe 2. Questionnaire UTAUT (Venkatesh et al., 2012a)

Attentes en termes d'Effort

- Je trouve [ce système] utile dans ma vie quotidienne
- Utiliser [ce système] m'aide à réaliser les choses plus rapidement
- Utiliser [ce système] augmente ma productivité

Attentes en termes de Performance

- Apprendre à utiliser [ce système] est facile pour moi
- Mon interaction avec [ce système] est claire et compréhensible
- Je trouve [ce système] facile à utiliser
- Il est facile pour moi de devenir compétent dans l'utilisation de [ce système]

Influence Sociale

- Les personnes qui sont importantes pour moi pensent que je devrais utiliser [ce système]
- Les personnes qui influencent mon comportement pensent que je devrais utiliser [ce système]
- Les personnes dont j'apprécie l'opinion préfèrent que j'utilise [ce système]

Conditions Facilitatrices

- J'ai les ressources nécessaires pour utiliser [ce système]
- J'ai les connaissances nécessaires pour utiliser [ce système]
- [Ce système] est compatible avec les autres technologies que j'utilise
- Je peux obtenir de l'aide lorsque j'ai des difficultés à utiliser [ce système]

Motivation Hédonique

- Utiliser [ce système] est fun
- Utiliser [ce système] est agréable
- Utiliser [ce système] est très divertissant

Valeur de Prix

- Le prix de [ce système] est raisonnable
- [Ce système] est un bon produit par rapport à son prix
- À ce prix, [ce système] a un bon rapport qualité/prix

Habitudes

- L'utilisation de [ce système] est devenue une habitude pour moi
- Je suis addict à l'utilisation de [ce système]
- Il faut que j'utilise [ce système]

Intention Comportementale

- J'ai l'intention de continuer à utiliser [ce système] dans le futur
- J'essaierai systématiquement d'utiliser [ce système] dans ma vie quotidienne
- Je prévois de continuer à utiliser [ce système] régulièrement

Annexe 3. Questionnaire Attrackdiff (Hassenzahl, 2004 ; traduit par Lallemand, Koenig, et al., 2015)

Qualités Hédoniques - Identification

- M’isole − Me sociabilise
- *Professionnel − Amateur
- *De bon goût − De mauvais goût
- Bas de gamme − Haut de gamme
- M’exclut − M’intègre
- *Me rapproche des autres − Me sépare des autres
- Non présentable – Présentable

Qualités Hédoniques – Stimulation

- *Original - Conventionnel
- Sans imagination - Créatif
- *Audacieux - Prudent
- *Novateur - Conservateur
- Ennuyeux - Captivant
- Peu exigeant - Challenging
- *Nouveau – Commun

Qualités Pragmatiques

- *Humain - Technique
- *Simple - Compliqué
- *Pratique - Pas pratique
- Fastidieux - Efficace
- *Prévisible - Imprévisible
- Confus - Clair
- Incontrôlable – Maîtrisable

Evaluation

- Laid − Beau
- *Bon - mauvais

* : items à renverser

Annexe 4. Questionnaire Avant Usage

Identifiant question	Items
Jugement global	
Quest1	Si vous deviez évaluer globalement cette application, quelle note lui donneriez-vous ?
Intention d'Usage	
Quest2	Je pense que je serais susceptible d'utiliser cette application dans le futur
Quest3	Il y a des chances que j'utilise cette application régulièrement
Quest4	La probabilité que j'utilise cette application est élevée
Utilité	
Quest5	J'estime que cette application pourrait répondre à mes besoins
Quest6	Je trouve que cette application pourrait m'être utile
Quest7	Je trouve que cette application offre des fonctionnalités qui pourraient m'être utiles
Quest8	Je pense qu'utiliser cette application pourrait être bénéfique pour moi
Utilisabilité	
Quest9	L'utilisation de cette application semble simple
Quest10	Je pense pouvoir comprendre rapidement comment utiliser cette application
Quest11	Se servir de cette application parait facile
Quest12	Il semble facile de comprendre comment utiliser cette application
Motivation	
	Consigne : Indiquez dans quelle mesure chacun des énoncés suivants correspond à l'une des raisons pour lesquelles vous pourriez utiliser cette application
Quest31	Parce que l'utilisation de cette application semble très fun
Quest32	Parce qu'utiliser cette application pourrait m'amuser beaucoup
Quest33	Parce que grâce à cette application je pourrais satisfaire ma curiosité à découvrir de nouvelles choses
Quest34	Parce que je pourrais avoir beaucoup de plaisir à interagir avec cette application
Quest35	Parce que cette application pourrait me permettre d'apprendre/de faire des choses que je trouve très intéressantes
Stimulation	
Quest13	Cette application semble originale
Quest14	Cette application me semble novatrice
Quest15	Les fonctions proposées par cette application paraissent innovantes
Quest16	Les fonctionnalités offertes par cette application semblent sortir du commun

Attractivité globale	
Quest17	L'utilisation de cette application semble plaisante
Quest18	Cette application semble agréable à utiliser
Quest18	Cette application paraît sympathique à utiliser
Influence Sociale	
Quest20	Les personnes dont j'apprécie l'opinion pourraient m'inciter à utiliser cette application
Quest21	Les personnes de mon entourage que j'ai l'habitude d'écouter pourraient me conseiller d'utiliser cette application
Quest22	Les personnes de mon entourage qui comptent pour moi pourraient m'encourager à utiliser cette application
Image de Soi	
Quest23	Les gens pourraient m'envier d'utiliser cette application
Quest24	Je pourrais être perçu(e) positivement par les autres si j'utilisais cette application
Quest25	L'utilisation de cette application renverrait de moi une image positive
Quest26	Utiliser cette application pourrait me valoriser auprès des autres
Confiance	
Quest27	Cette application semble fournir des informations fiables
Quest28	J'ai confiance dans les informations que cette application pourrait fournir
Quest29	Cette application semble fiable
Quest30	Je pense que l'on peut avoir confiance en cette application

Annexe 5. Questionnaire Après Usage

Identifiant question	Items
Jugement global	
Quest56	Si vous deviez évaluer globalement cette application, quelle note lui donneriez-vous ?
Quest58	À quel point recommanderiez-vous cette application à votre entourage ?
Intention d'Usage	
Quest1	J'ai l'intention de continuer à utiliser cette application
Quest2	Il y a de fortes chances pour que j'utilise cette application dans ma vie quotidienne
Quest3	Je prévois de continuer à utiliser cette application régulièrement
Quest4	Je pense poursuivre mon utilisation de cette application
Utilité	
Quest5	J'estime que cette application répond à mes besoins
Quest6	Je trouve que cette application m'est utile
Quest7	Je trouve que cette application offre des fonctionnalités qui me seront utiles
Quest8	Je pense qu'utiliser cette application est bénéfique pour moi
Utilisabilité	
Quest9	L'utilisation de cette application est simple
Quest10	J'ai rapidement compris comment utiliser cette application
Quest11	Se servir de cette application est facile
Quest12	Il est facile de comprendre comment utiliser cette application

Motivation	
	Consigne : Indiquez dans quelle mesure chacun des énoncés suivants correspond à l'une des raisons pour lesquelles vous utilisez cette application
Quest51	• Parce que l'utilisation de cette application est très fun
Quest52	• Parce qu'utiliser cette application m'amuse beaucoup
Quest53	• Parce que grâce à cette application je peux satisfaire ma curiosité à découvrir de nouvelles choses
Quest54	• Parce que j'ai beaucoup de plaisir à interagir avec cette application
Quest55	• Parce que cette application me permet d'apprendre/de faire des choses que je trouve très intéressantes
Stimulation	
Quest13	Je trouve cette application originale
Quest14	Je trouve cette application novatrice
Quest15	Les fonctions proposées par cette application sont innovantes
Quest16	Les fonctionnalités offertes par cette application sortent du commun
Attractivité globale	
Quest17	L'utilisation de cette application est plaisante
Quest18	Cette application est agréable à utiliser
Quest19	Cette application est sympathique à utiliser
Influence Sociale	
Quest27	Les personnes dont j'apprécie l'opinion m'incitent à utiliser cette application
Quest28	Les personnes de mon entourage que j'ai l'habitude d'écouter me conseillent d'utiliser cette application
Quest29	Les personnes de mon entourage qui comptent pour moi m'encouragent à utiliser cette application
Image de Soi	
Quest30	Les gens m'envient d'utiliser cette application
Quest31	Je suis perçu(e) positivement par les autres quand j'utilise cette application
Quest32	L'utilisation de cette application renvoie une image positive de moi
Quest33	Utiliser cette application valorise mon image auprès des autres

Confiance	
Quest24	Je trouve que cette application fournit des informations fiables
Quest25	J'ai confiance dans les informations fournies par cette application
Quest26	Les informations apportées par cette application semblent de bonne qualité
Esthétisme	
Quest20	J'aime l'apparence visuelle de cette application
Quest21	La présentation de cette application est agréable à regarder
Quest22	Cette application est esthétiquement attrayante
Quest23	Cette application est belle
Absorption	
Quest34	Je me sens à certains moments complètement absorbé(e) par l'interaction avec ce système
Quest35	Il m'arrive de perdre la notion du temps quand j'utilise ce système
Quest36	Je ne vois pas le temps passer lorsque j'utilise ce système
Quest37	Quand j'utilise cette application, je ne pense plus à rien d'autre
Quest38	Quand j'interagis avec cette application, j'oublie tout le reste
Emotions	
	Consigne : quand j'utilise cette application, je me sens
Quest39	• Découragé(e), abattu(e), déprimé(e)
Quest40	• Contrarié(e), mécontent(e)
Quest41	• Nerveux(se), agité(e), impatient(e)
Quest42	• Agacé(e), énervé(e), exaspéré(e)
Quest43	• Frustré(e), déçu(e), insatisfait(e)
Quest44	• Mal
Quest45	• Intéressé(e)
Quest46	• Enthousiaste, excité(e)
Quest47	• Joyeux(se), gai(e)
Quest48	• Content(e), satisfait(e)
Quest49	• Serein(e), calme
Quest50	• Bien

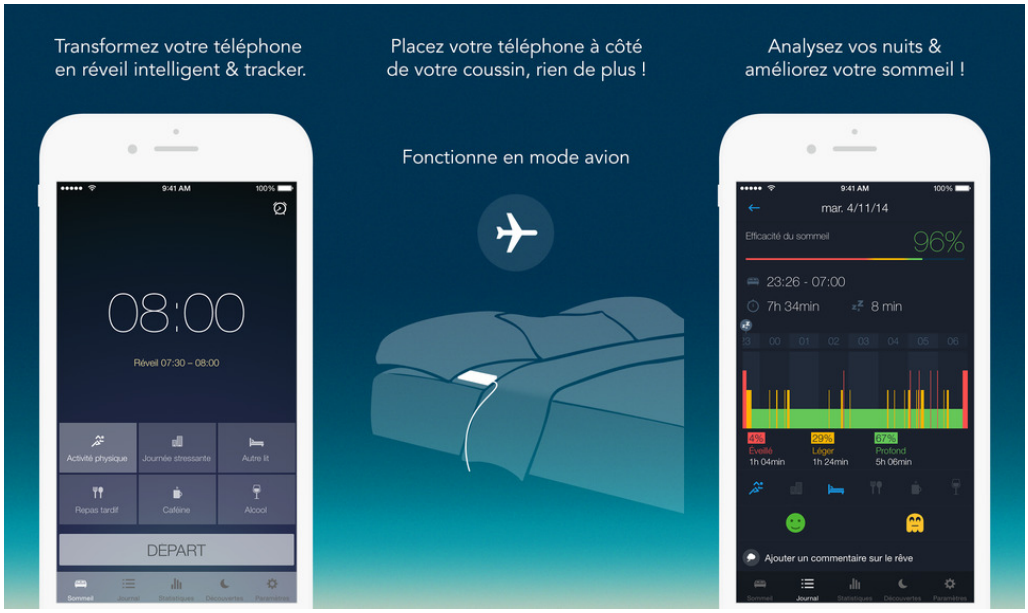
Habitudes	
Quest51	Utiliser cette application est devenu automatique pour moi
Quest52	J’ai l'habitude d"utiliser cette application
Quest53	Quand la situation s'y prête, utiliser cette application est un choix évident pour moi
Quest54	Utiliser cette application est naturel pour moi

Annexe 6. Modèles en équations structurelles (SEM)

Lorsqu'un nouveau modèle théorique est conceptualisé, il apparait nécessaire d'en faire une analyse rigoureuse afin de tester les hypothèses posées à partir de la littérature. Généralement, les modèles sont, au moins en partie, composés de liens supposés entre des variables. L'examen du modèle vise alors généralement à tester l'existence de ces liens. Pour cela, plusieurs techniques sont utilisées dans la littérature : l'analyse de régression et la modélisation en équations structurelles (*Structural Equation Modelling* – SEM). La modélisation SEM peut être considérée comme un cadre générale d'analyse dont la régression est un cas particulier (Fan, 2003).

Les SEMs (se référer à Hoyle (2012) pour une revue complète) présentent deux avantages majeurs par rapport aux analyses de régressions. Premièrement, les modèles de régressions permettent uniquement d'évaluer l'influence d'une variable (ou plusieurs variables) sur une autre, alors que les SEMs offrent la possibilité d'analyser un modèle dans son ensemble (Law & van Schaik, 2010). Deuxièmement, les SEMs (Gefen, Straub, & Boudreau, 2000), contrairement aux régressions, offrent la possibilité d'évaluer à la fois le modèle structurel (i.e., liens entre les variables) mais aussi le modèle de mesure (i.e., liens entre variables observables et variables latentes). Le modèle structurel correspond à la spécification des liens entre les variables du modèle. Quant au modèle de mesure, il explicite les liens entre les variables observables et les variables latentes (voir Chapitre 4 : Construction des échelles à items uniques pour la mesure avant et après usage). Enfin, les paramètres, dans le cadre des SEMs, sont estimés de manière simultanée (Lei & Wu, 2007). À l'inverse, dans le cadre des régressions, les équations sont estimées une par une (Alavifar, Karimimalayer, & Anuar, 2012). Les SEMs offrent, par conséquent, la possibilité de vérifier à partir de jeux de données que les liens supposés entre des variables sont valides (Juhel, 2015). Il est également possible de tester l'ajustement du modèle (i.e., liens entre les variables) aux données.

Annexe 7. Description de l'application Sleep Better



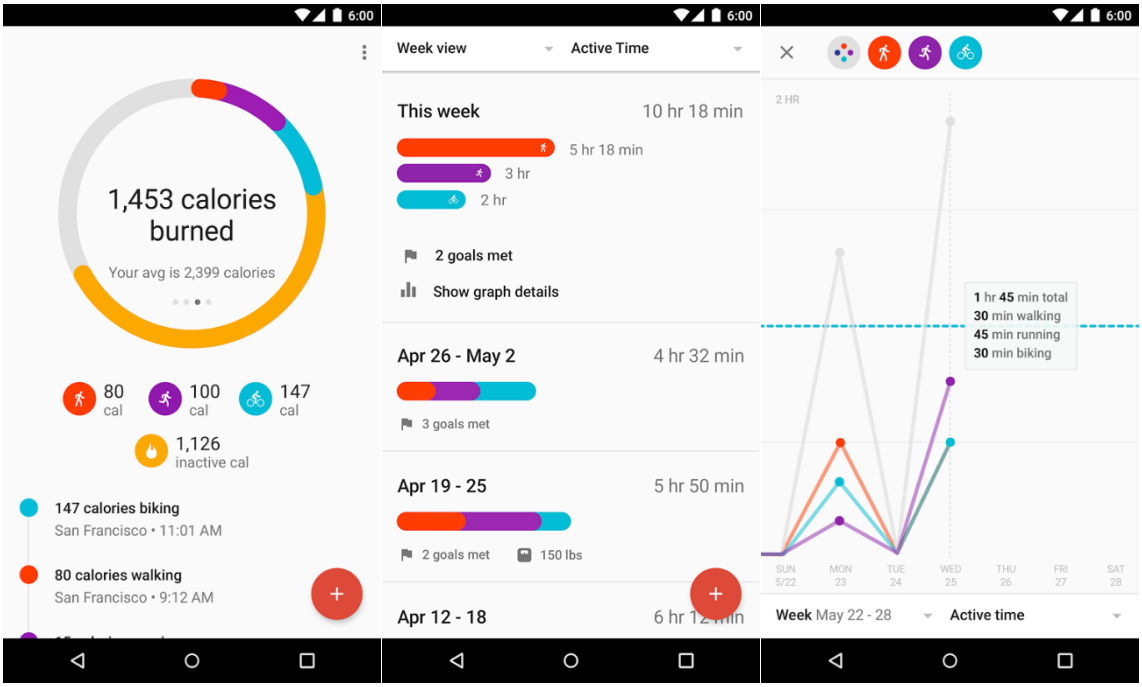


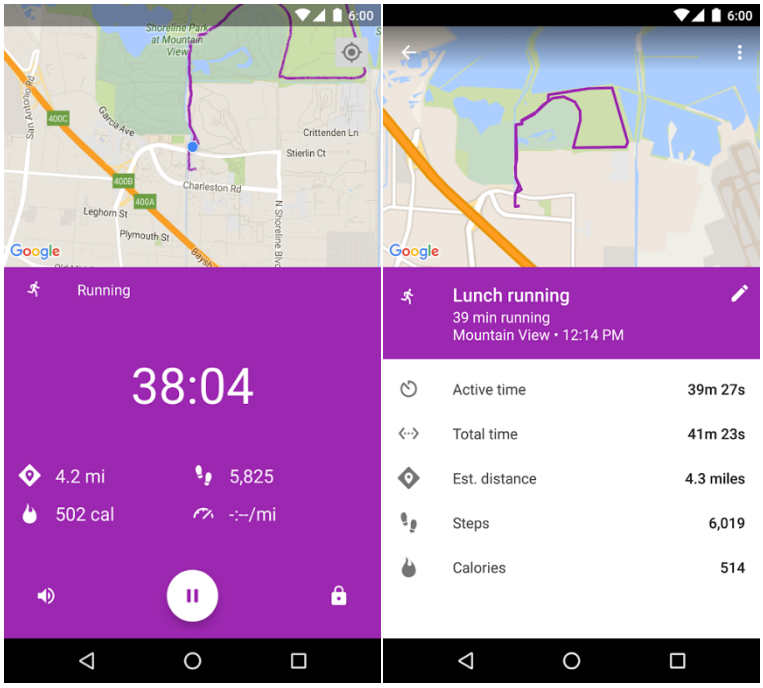
Suivez votre sommeil, analysez vos rêves, améliorez vos heures de coucher et réveillez-vous mieux que jamais avec Sleep Better de Runtastic ! Vous avez des difficultés à vous endormir ? Vous voulez découvrir comment vos activités quotidiennes influencent l'efficacité de votre nuit ? Ou vous voulez tout simplement de nouvelles sonneries de réveil pour vous aider à mieux attaquer la journée ? Sleep Better vous permet d'améliorer la qualité de votre sommeil de façon simple et captivante !

Obtenez un aperçu de votre nuit, distinguez les phases lunaires, écrivez un journal de rêves et bien plus encore grâce à votre smartphone :

- Enregistrez votre durée, vos cycles et votre efficacité de sommeil à l'aide de simples clics
- Profitez de toutes les fonctions de suivi, même en Mode avion, en posant tout simplement votre appareil sur votre lit (à côté de votre coussin)
- Discernez entre sommeil léger, sommeil profond & temps éveillé au lit (avant de vous endormir et en vous réveillant)
- Utilisez la Smart Alarm pour être réveillé au meilleur moment, dans l'espace de temps que vous aurez choisi
- Profitez de nombreuses sonneries de réveil et de la fonction snooze pour débuter parfaitement la journée
- Entrez des habitudes journalières & des facteurs, tels que la consommation de caféine, d'alcool, l'activité physique et le niveau de stress, afin de voir leur influence sur la qualité de votre sommeil
- Observez les phases lunaires
- Gardez un journal de rêves pour faire un constat de ce que vous rêvez le plus et connaître le pourcentage de bons rêves, rêves neutres ou mauvais rêves
- Commentez votre humeur du matin pour avoir des statistiques supplémentaires
- Voir l'historique - votre suivi personnel - via un design épuré et moderne, incluant des statistiques et des aperçus détaillés pour des périodes plus longues (par ex. semaine ou mois spécifique)
- Filtrez l'historique en prenant en compte des facteurs journaliers pour déterminer lorsque vous dormez le mieux
- Partagez vos sessions de sommeil via Google+, Facebook, Twitter, WhatsApp & email
- Fonctionne sous iOS et Android

Annexe 8. Description de l'application Google Fit





- Effectuez le suivi de votre activité en toute simplicité. Lorsque vous marchez, courez ou pédalez dans la journée, votre téléphone ou votre montre Android Wear enregistrent automatiquement les données associées à votre activité avec Google Fit.
- Obtenez des informations immédiates. Découvrez en temps réel les statistiques liées à vos sessions de course, à vos promenades et à vos sorties à vélo.

Google Fit enregistre votre vitesse, votre rythme, l'itinéraire suivi, l'altitude et d'autres informations pour vous aider à rester motivé et à atteindre vos objectifs.

- Atteignez vos objectifs de remise en forme. Fixez-vous des objectifs de nombre de pas, de durée, de distance ou de calories brûlées. Recevez des recommandations personnalisées et des encouragements pour atteindre vos objectifs.
- Évaluez votre bien-être général. Accédez aux informations d'autres applications de suivi de remise en forme, de nutrition, de sommeil et de poids : Android Wear, Nike+, Runkeeper, Strava, MyFitnessPal, Lifesum, Basis, Sleep as Android, Withings et Xiaomi Mi bands.
- Utilisez l'application où que vous soyez. Suivez vos progrès sur votre téléphone ou votre tablette, sur le Web (fit.google.com), et même sur votre montre Android Wear.

Annexe 9. Statistiques descriptives des données collectées avec les échelles avant usage durant l'étude Sleep Better

Tableau 27. Statistiques descriptives des items de l'échelle A-M

Variable latente	**Motivation Intrinsèque**					**Image de Soi**			
Item	IM1	IM2	IM3	IM4	IM5	SEI1	SEI2	SEI3	SEI4
Moyenne	4.75	4.89	5.80	4.61	5.43	3.26	3.09	3.36	2.61
E.-T.	2.60	2.67	2.71	2.61	2.78	2.62	2.39	2.43	2.29

Variable latente	**Influence Sociale**		
Item	SOI1	SOI2	SOI3
Moyenne	4.14	3.97	3.81
E.-T.	2.84	2.75	2.90

Tableau 28. Statistiques descriptives des items de l'échelle UX

Variable latente	**Utilité**				**Utilisabilité**			
Item	USE1	USE2	USE3	USE4	USA1	USA2	USA3	USA4
Moyenne	4.95	5.44	5.81	5.33	6.98	7.53	6.95	7.15
E.-T.	2.97	3.09	2.99	2.89	1.94	1.90	2.03	1.92

Variable latente	**Stimulation**				**Confiance**			
Item	ST1	ST2	ST3	ST4	TR1	TR2	TR3	TR4
Moyenne	6.62	6.35	6.39	6.25	4.96	4.60	5.02	4.83
E.-T.	2.28	2.49	2.41	2.33	2.25	2.36	2.25	2.45

Tableau 29. Statistiques descriptives des variables observables de l'acceptabilité

Variable latente	**Intention d'Usage**			**Jugement Global**
Item	ITU1	ITU2	ITU3	OJ
Moyenne	5.28	4.56	4.67	6.09
E.-T.	3.20	3.08	3.11	2.00

Annexe 10. Statistiques descriptives des données collectées avec les échelles avant usage durant l'étude Google Fit

Tableau 30. Statistiques descriptives des items de l'échelle A-M

Variable latente	**Motivation Intrinsèque**					**Image de Soi**			
Item	IM1	IM2	IM3	IM4	IM5	SEI1	SEI2	SEI3	SEI4
Moyenne	3.60	4.22	4.38	4.15	4.24	2.78	3.42	4.06	3.02
E.-T.	2.65	2.71	2.67	2.67	2.77	2.45	2.67	2.91	2.66

Variable latente	**Influence Sociale**		
Item	SOI1	SOI2	SOI3
Moyenne	3.94	3.82	3.60
E.-T.	2.89	2.84	2.72

Tableau 31. Statistiques descriptives des items de l'échelle UX

Variable latente	**Utilité**				**Utilisabilité**			
Item	USE1	USE2	USE3	USE4	USA1	USA2	USA3	USA4
Moyenne	4.79	4.93	5.39	5.30	7.06	7.65	7.06	7.17
E.-T.	2.63	2.74	2.67	2.56	1.82	1.94	1.94	1.91

Variable latente	**Stimulation**				**Confiance**			
Item	ST1	ST2	ST3	ST4	TR1	TR2	TR3	TR4
Moyenne	4.09	4.10	4.21	3.68	5.85	5.39	5.89	5.46
E.-T.	2.36	2.33	2.32	2.23	2.09	2.24	2.20	2.32

Tableau 32. Statistiques descriptives des variables observables de l'acceptabilité

Variable latente	**Intention d'Usage**			**Jugement Global**
Item	ITU1	ITU2	ITU3	OJ
Moyenne	4.61	3.91	3.91	5.95
E.-T.	2.92	2.71	2.81	1.84

Annexe 11. Analyses factorielles exploratoires sur les données collectées avec les échelles avant usage durant l'étude Sleep Better

Tableau 33. Résultats de l'AFE pour l'échelle UX[61]

Items	Facteurs			
	Utilité	Utilisabilité	Stimulation	Confiance
USE1	.91			
USE2	.93			
USE3	.91			
USE4	.92			
USA1		.90		
USA2		.70		
USA3		.86		
USA4		.91		
STI1			.88	
STI2			.88	
STI3			.86	
STI4			.90	
TRU1				.94
TRU2				.90
TRU3				.88
TRU4				.65
Variance expliquée	*.22*	*.18*	*.20*	*.19*

[61] Seules les saturations supérieures à .30 sont présentées.

Tableau 34. Résultats de l'AFE pour l'échelle A-M[61]

Items	**Facteurs**		
	Motivation Intrinsèque	Image de Soi	Influence Sociale
IMO1	.82		
IMO2	.86		
IMO3	.79		
IMO4	.81		
IMO5	.73		
SEI1		.64	
SEI2		.93	
SEI3		.69	
SEI4		.80	
SOI1			.87
SOI2			.92
SOI3			.82
Variance expliquée	*.27*	*.21*	*.21*

Tableau 35. Résultats AFE pour l'Intention d'Usage[61]

Items	**Facteur**
	Intention d'Usage
INT1	.95
INT2	.94
INT3	.97
Variance expliquée	*.90*

Annexe 12. Statistiques descriptives des données collectées avec les échelles après usage durant l'étude Waze

Tableau 36. Statistiques descriptives des items de l'échelle A-M

Variable latente	**Motivation Intrinsèque**					**Image de Soi**			
Item	IM1	IM2	IM3	IM4	IM5	SEI1	SEI2	SEI3	SEI4
Moyenne	3.60	4.22	4.38	4.15	4.24	2.78	3.42	4.06	3.02
E.-T.	2.65	2.71	2.67	2.67	2.77	2.45	2.67	2.91	2.66

Variable latente	**Influence Sociale**		
Item	SOI1	SOI2	SOI3
Moyenne	3.94	3.82	3.60
E.-T.	2.89	2.84	2.72

Variable latente	**Habitudes**				**Emotions Négatives**					
Item	IM1	IM2	IM3	IM4	EMN1	EMN2	EMN3	EMN4	EMN5	EMN6
Moyenne	3.60	4.22	4.38	4.15	0.76	1.10	1.27	1.33	1.07	0.95
E.-T.	2.65	2.71	2.67	2.67	1.73	1.91	1.96	2.28	1.97	1.82

Variable latente	**Emotions Positives**						**Absorption**				
Item	EMP1	EMP2	EMP3	EMP4	EMP5	EMP6	ABS1	ABS2	ABS3	ABS4	ABS5
Moyenne	4.38	2.90	3.40	5.64	5.47	5.06	2.10	1.36	2.44	1.53	1.52
E.-T.	2.69	2.62	2.61	2.85	2.62	2.62	2.49	2.02	2.61	2.28	2.16

Tableau 37. Statistiques descriptives des items de l'échelle UX

Variable latente	**Utilité**			**Utilisabilité**			
Item	USE1	USE2	USE3	USA1	USA2	USA3	USA4
Moyenne	8.51	8.09	6.80	7.78	7.95	7.96	7.89
E.-T.	1.71	1.90	2.73	2.14	2.22	2.14	2.09

Variable latente	**Stimulation**				**Confiance**			**Esthétisme**			
Item	ST1	ST2	ST3	ST4	TR1	TR2	TR3	AES1	AES2	AES3	AES4
Moyenne	5.87	6.47	6.07	5.46	7.83	7.78	7.83	6.76	6.59	6.15	5.90
E.-T.	2.69	2.47	2.58	2.75	1.78	1.60	1.61	2.27	2.42	2.49	2.33

Tableau 38. Statistiques descriptives des variables observables de l'acceptabilité

Variable latente	**Intention d'Usage**			
Item	ITU1	ITU2	ITU3	ITU4
Moyenne	8.55	7.44	8.01	8.37
S.D.	2.18	2.70	2.53	2.34

Annexe 13. Scree Plot sur les données récoltées avec les échelles UX et A-M après usage

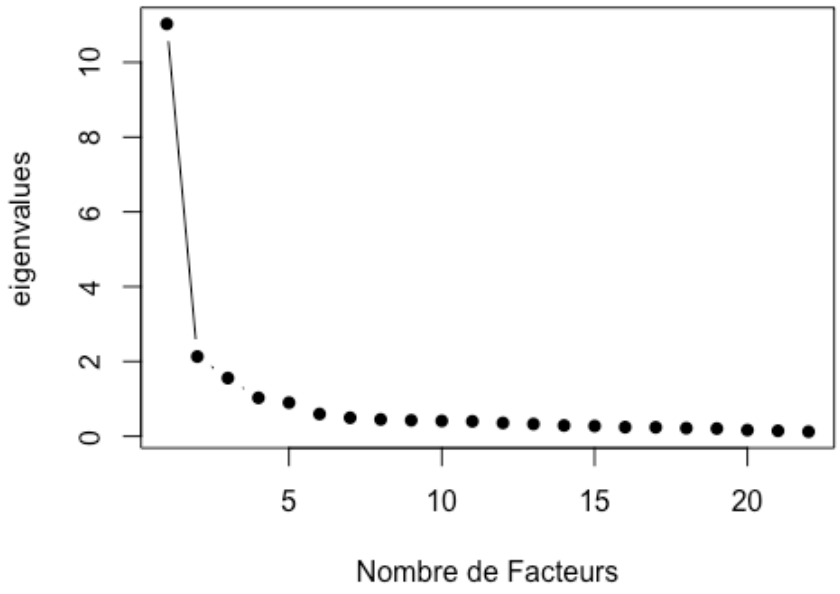


Figure 44. Scree Plot sur les données récoltées avec l'échelle UX après usage

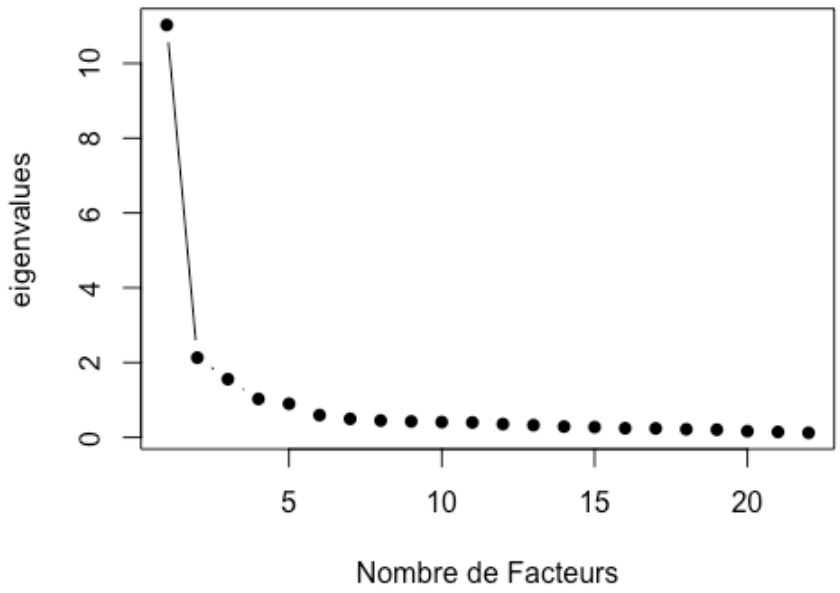


Figure 45. Scree Plot sur les données récoltées avec l'échelle A-M après usage

Annexe 14. Analyses factorielles exploratoires sur les données collectées avec les échelles après usage durant l'étude Waze

Tableau 39. Résultats des AFE pour l'échelle UX[62]

Items	Facteurs				
	Utilité perçue	Utilisabilité perçue	Stimulation	Confiance	Esthétisme
USE1	*.35*	*.37*			
USE2	.70				
USE3	.53				
USE4	.64				
USA1		.82			
USA2		.88			
USA3		.88			
USA4		.90			
STI1			.70		
STI2			.83		
STI3			.90		
STI4			.74		
TRU1				.95	
TRU2				.69	
TRU3				.78	
AES1					.86
AES2					.77
AES3					.69
AES4					.83

[62] Seules les saturations supérieures à .30 sont présentées. De plus, les items présentant des saturations sur plus d'un facteur sont présentés en italiques.

Tableau 40. Résultats des AFE pour l'échelle A-M[63]

Items	Facteurs						
	Influence Sociale	Image de Soi	Absorption	Émotions Négatives	Émotions Positives	Motivation Intrinsèque	Habitudes
SOI1	.79						
SOI2	.83						
SOI3	.85						
SEI1		.39					
SEI2		.74					
SEI3		.68					
SEI4		.81					
ABS1			.64				
ABS2			.49				
ABS3			*.52*			*.34*	
ABS4			.76				
ABS5			.66				
EMN1				.79			
EMN2				.88			
EMN3				.56			
EMN4				.84			
EMN5				.61			
EMN6				.91			
EMP1					*.39*	*.54*	
EMP2					*.32*	*.36*	
EMP3					.69		
EMP4					.57		
EMP5					.49		
EMP6					.68		
IMO1						.49	
IMO2						.49	
IMO3						.66	
IMO4					*.31*		
IMO5						.87	
HAB1							.96
HAB2							.89
HAB3							.71
HAB4							.84

[63] Seules les saturations supérieures à .30 sont présentées. De plus, les items présentant des saturations sur plus d'un facteur sont présentés en italiques.

Annexe 15. Résultats analyses factorielles confirmatoires pour l'échelle A-M

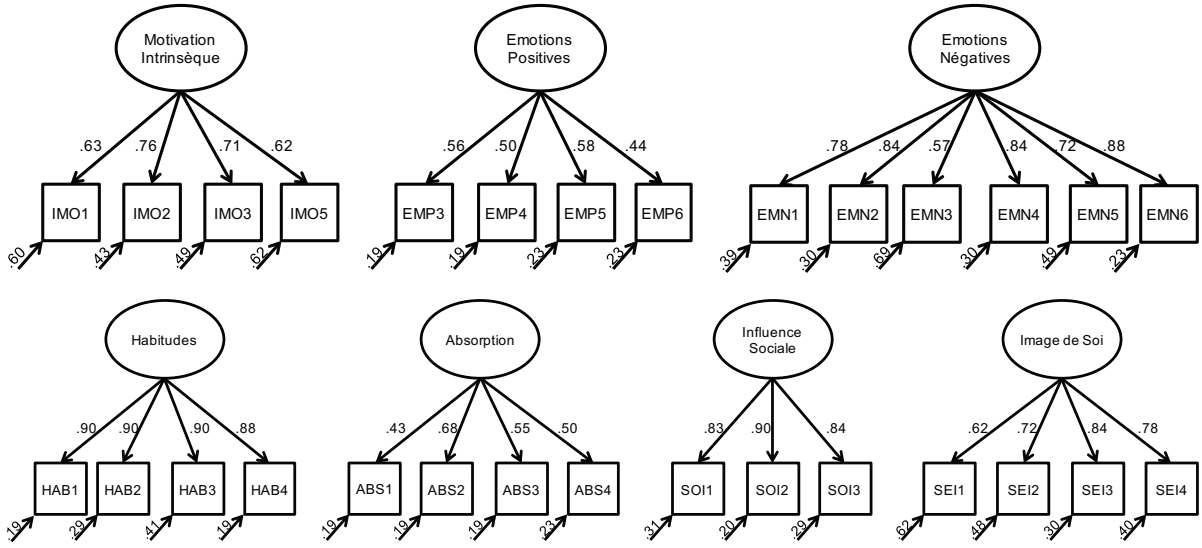


Figure 46. Résultats analyses factorielles confirmatoires pour l'échelle A-M après usage[64]

[64] Le nombre de covariances entre les variables étant trop nombreuses, les résultats sont rapportés dans le Tableau 41.

Tableau 41. Covariances entre les variables latentes pour l'échelle A-M après usage

Covariance		Estimation complètement standardisée
Influence Sociale	~~	
	Image de Soi	0,50
	Motivation	0,28
	Habitudes	0,38
	Emotions Négatives	0,08
	Emotions Positives	0,43
	Absorption	0,32
Image de Soi	~~	
	Motivation	0,68
	Habitudes	0,18
	Emotions Négatives	0,22
	Emotions Positives	0,54
	Absorption	0,56
Motivation		
	Habitudes	0,22
	Emotions Négatives	0,09
	Emotions Positives	0,63
	Absorption	0,64
Habitudes	~~	
	Emotions Négatives	-0,40
	Emotions Positives	0,58
	Absorption	0,21
Emotions Négatives	~~	
	Emotions Positives	-0,09
	Absorption	0,31
Emotions Positives	~~	
	Absorption	0,53

Annexe 16. Autres modélisations sur les données après usage

Au sein du Chapitre 2 : Jugement après usage, des modèles structuraux basés sur la covariance (*Covariance Based-Structural Equation Modeling ; CB-SEM*) ont été employés pour modéliser le jugement après usage. Cette même approche a été utilisée tout au long de la thèse pour évaluer simultanément les modèles de mesure (i.e., structures factorielles des outils de mesure) ainsi que les modèles structuraux (i.e., modèle théorique supposant les liens entre les variables). Cependant, cette méthode requière des conditions d'utilisation parfois difficile à atteindre comme le nombre de participants ou encore le nombre de variables observables par variable latente.

Dans le cadre de l'expérimentation sur l'après usage, les résultats ont montré plusieurs limites. En effet, des effets bien étayés dans la littérature (e.g., effet de l'utilisabilité perçue sur l'intention d'usage ; King & He, 2006) ne sont pas retrouvés dans les résultats de notre étude. Ces différences pourraient s'expliquer par la méthode de modélisation utilisée (i.e., SEM). Dans la littérature, deux méthodes majoritaires sont présentes avec d'une part les modèles structuraux basés sur la covariance (CB-SEM) et d'autre part les modèles structuraux reposant sur les moindres carrés partiels (*Partial Least Square-Structural Equation Modeling* ; PLS-SEM). Les différences entre ces méthodes se retrouvent dans l'estimation des paramètres. En effet, dans l'approche CB-SEM, les paramètres sont estimés de sorte à reconstruire la matrice théorique de variances/covariances par maximum de vraisemblance[65] alors que dans l'approche PLS-SEM les paramètres sont estimés de sorte à maximiser la variance expliquée des variables dépendantes du modèle (Amaro, Abrantes, & Seabra, 2015). De plus, des régressions multiples peuvent être utilisées (Gefen et al., 2000) pour tester les effets de plusieurs variables sur une variable dépendante.

Afin de tester les éventuelles différences avec l'approche CB-SEM, une régression linéaire multiple ainsi qu'un modèle PLS-SEM ont été estimés. Comme lors

[65] Des estimateurs « robustes » peuvent être employés lorsque les conditions de normalité multivariée ne sont pas respectées.

de l'étude présentée dans le Chapitre 2 : Jugement après usage, la variable Habitudes a été supprimée des modélisations à cause de la corrélation très élevée entre l'Intention d'Usage et les Habitudes (r = .97, p = .001).

En ce qui concerne la régression linéaire multiple, des scores moyens par variable non observable (e.g., Influence Sociale) ont été calculés avant l'estimation des paramètres. Les résultats sont présentés dans le Tableau 42. En résumé, les variables Utilité, Utilisabilité, Confiance, Influence Sociale et Emotions Négatives auraient une influence significative sur l'Intention de continuer à utiliser l'application mobile.

Tableau 42. Résultats de la régression linéaire multiple sur les données après usage[66]

Variable dépendante	Variables indépendantes	Bêta
Intention d'usage	Utilité	.33 ***
R^2 = .69	Utilisabilité	.32 ***
	Stimulation	-.03
	Confiance	.19 *
	Esthétisme	.07
	Influence Sociale	.12 ***
	Image de Soi	.00
	Motivation Intrinsèque	-.04
	Emotions Négatives	-.26 ***
	Emotions Positives	.09
	Absorption	.00

Concernant la modélisation PLS-SEM, la même configuration que pour le CB-SEM est utilisée pour le modèle de mesure et du modèle structurel. Les résultats sont présentés dans le Tableau 43[67]. En résumé, les variables Utilité, Utilisabilité, Influence Sociale et Emotions Négatives ont une influence significative sur l'Intention Comportementale envers l'application mobile.

[66] * = p < .05 ; ***: p < .001.

[67] Afin d'estimer la significativité des paramètres, une méthode bootstrap (*adjusted bootstrap percentile method*) est employée avec 500 échantillonnages.

Tableau 43. Résultats de modélisation PLS-SEM sur les données après usage[66]

Variable dépendante	Variables indépendantes	Coefficients bêta (IC 95%)
Intention d'usage	Utilité	[.21 ; .49] *
R^2 = .71	Utilisabilité	[.12 ; .47] *
	Stimulation	[-.11 ; .13]
	Confiance	[-.04 ; .21]
	Esthétisme	[-.08 ; .16]
	Influence Sociale	[.04 ; .25] *
	Image de Soi	[-.12 ; .07]
	Motivation	[-.09 ; .12]
	Emotions Négatives	[-.25 ; -.02] *
	Emotions Positives	[-.02 ; .18]
	Absorption	[-.03 ; .14]

Annexe 17. Analyses factorielles *bootstrapped* pour l'échelle UX et A-M avant usage

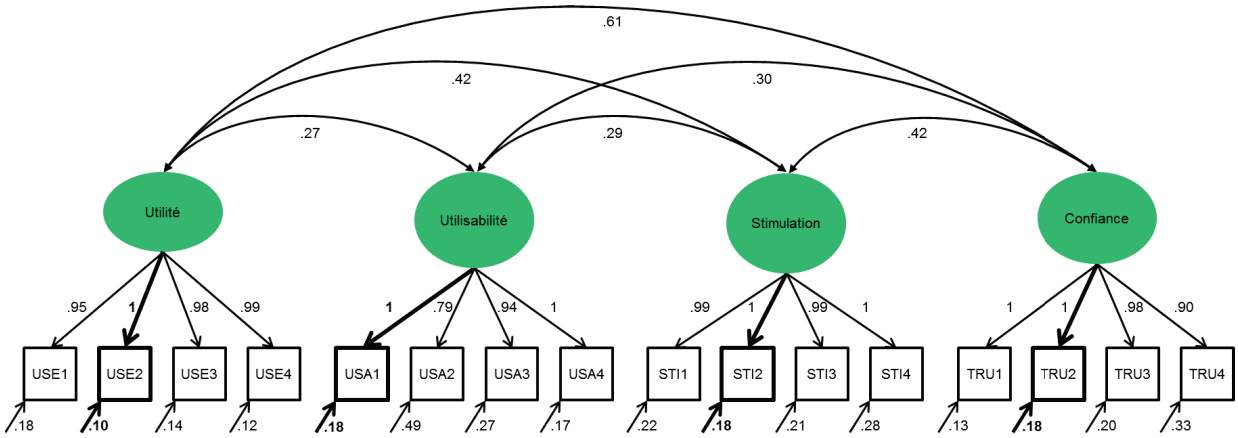


Figure 47. Résultats de l'analyse factorielle bootstrapped pour l'échelle UX[68]

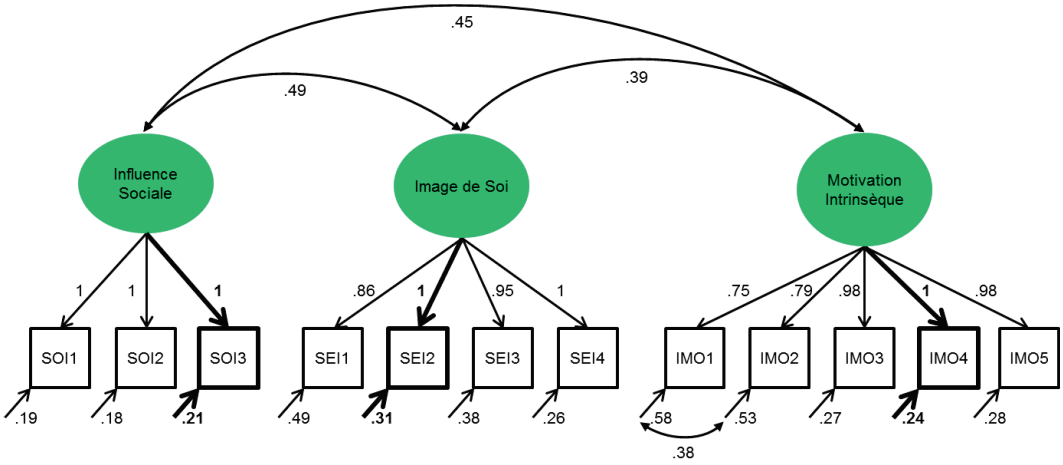


Figure 48. Résultats de l'analyse factorielle bootstrapped pour l'échelle A-M[68]

[68] Les contours en gras indiquent les items sélectionnés pour les échelles à items uniques.

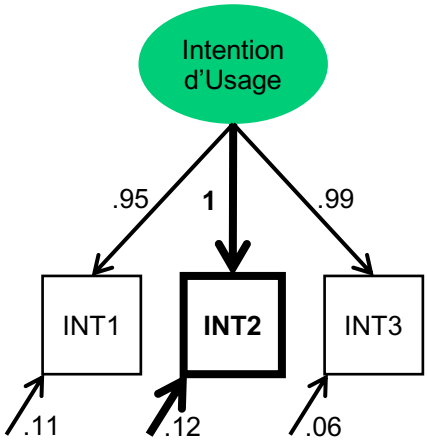


Figure 49. Résultats de l'analyse factorielle bootstrapped pour l'Intention d'Usage[68]

Annexe 18. Résultats des régressions logistiques

Utilité perçue

β = .57, p < .001

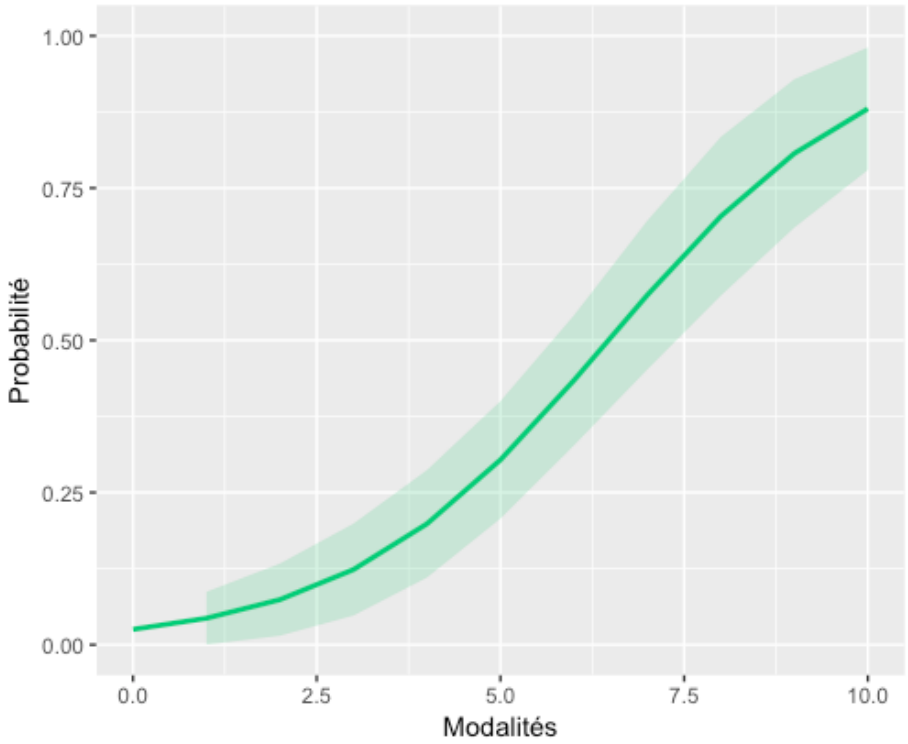


Figure 50. Régression logistique : probabilité d'adoption selon le niveau d'Utilité perçue

Tableau 44. Probabilité d'adoption en fonction du score d'Utilité perçue

Modalité Utilité	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.03	.04	.07	.12	.20	.30	.43	.57	.70	.81	.88

Utilisabilité perçue

β =.02, p = .80

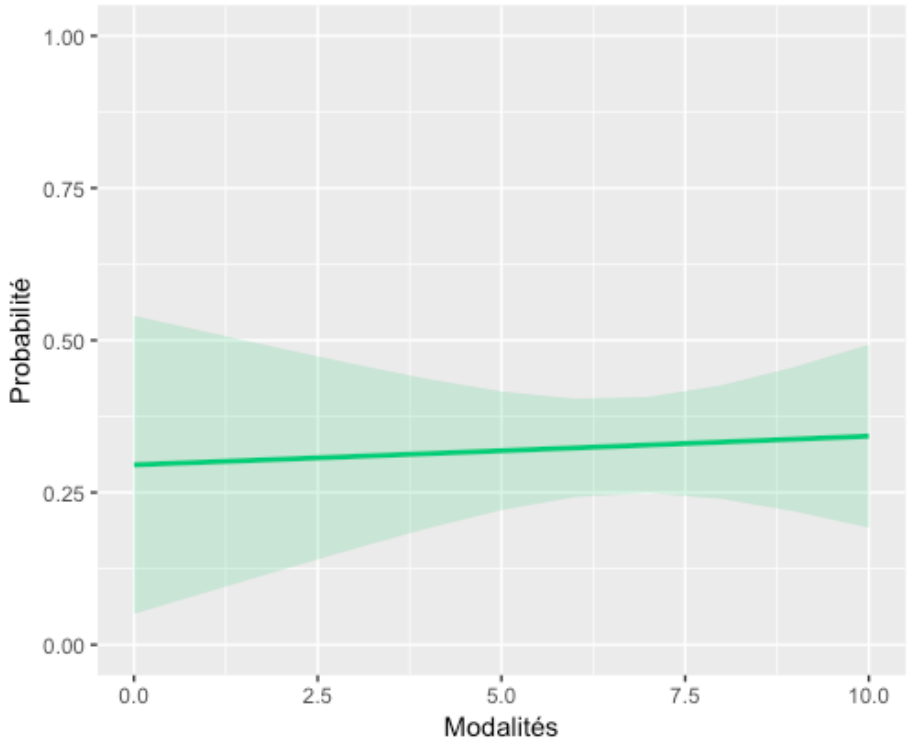


Figure 51. Régression logistique : probabilité d'adoption selon le niveau d'Utilisabilité perçue

Tableau 45. Probabilité d'adoption en fonction du score d'Utilisabilité perçue

Modalité Utilisabilité	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.30	.30	.30	.31	.31	.32	.32	.33	.34	.34	.34

Stimulation perçue

β =.14, p = .08

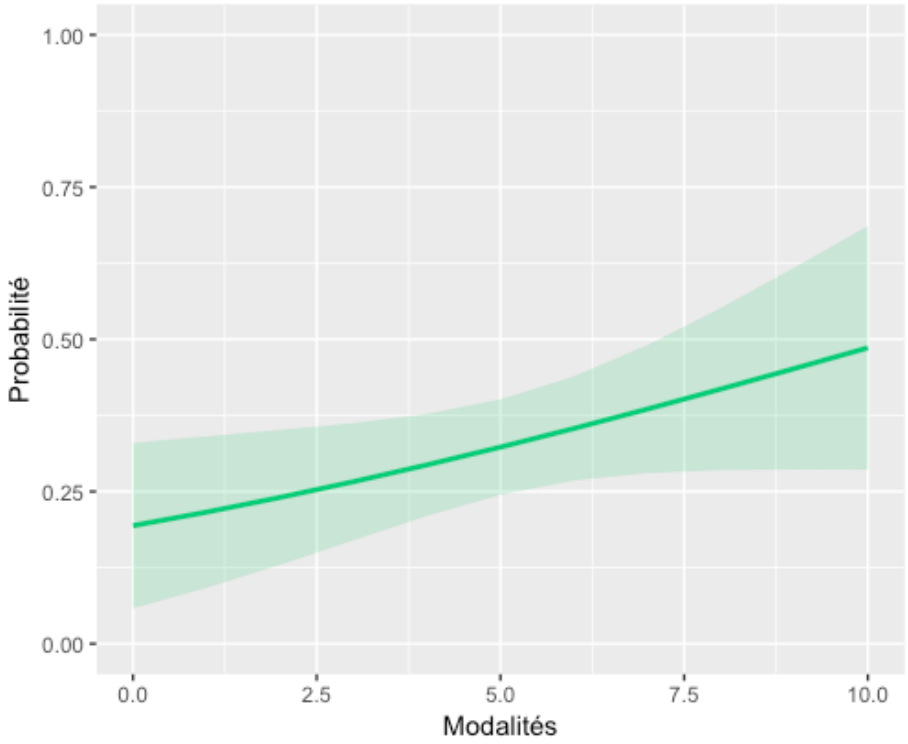


Figure 52. Régression logistique : probabilité d'adoption selon le niveau de Stimulation perçue

Tableau 46. Probabilité d'adoption en fonction du score de Stimulation perçue

Modalité Stimulation	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.19	.22	.24	.27	.29	.32	.35	.39	.42	.45	.49

Confiance perçue

β =.15, p = .09

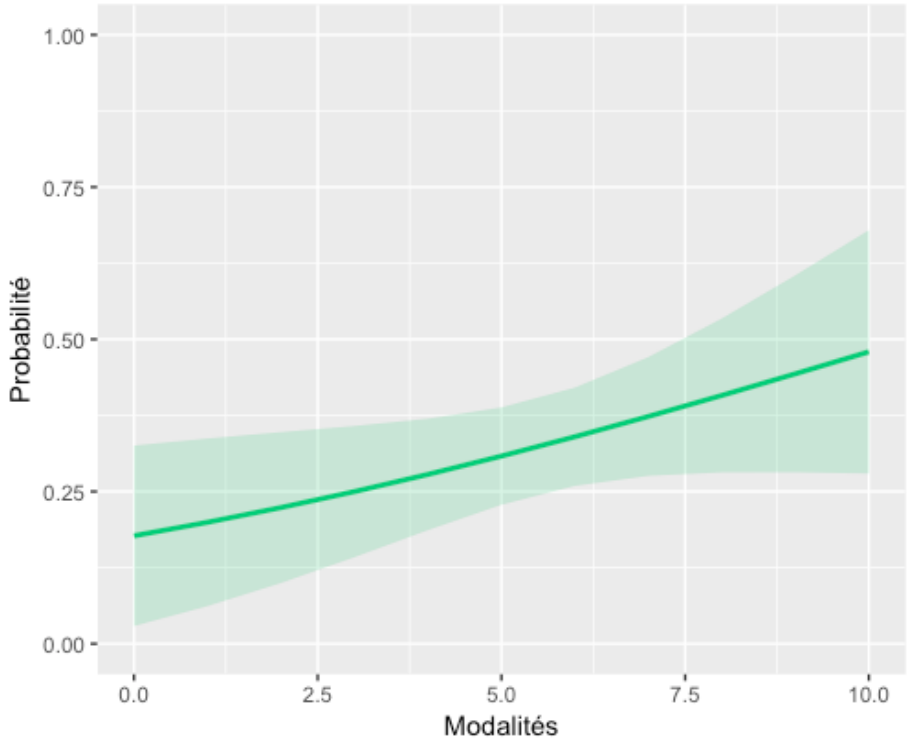


Figure 53. Régression logistique : probabilité d'adoption selon le niveau de Confiance perçue

Tableau 47. Probabilité d'adoption en fonction du score de Confiance perçue

Modalité Confiance	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.18	.20	.22	.25	.28	.30	.34	.37	.41	.44	.48

Influence Sociale

β =.26, p < .001

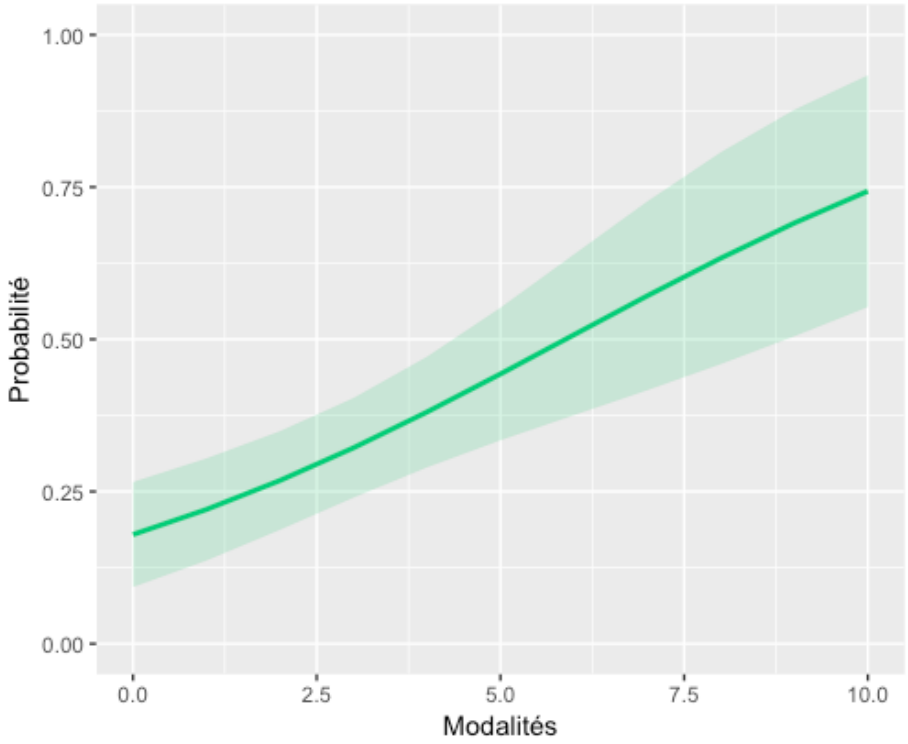


Figure 54. Régression logistique : probabilité d'adoption selon le niveau d'Influence Sociale

Tableau 48. Probabilité d'adoption en fonction du score d'Influence Sociale

Modalité Influence Sociale	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.18	.22	.27	.32	.38	.44	.51	.58	.63	.69	.74

Image de Soi

β =.17, p < .001

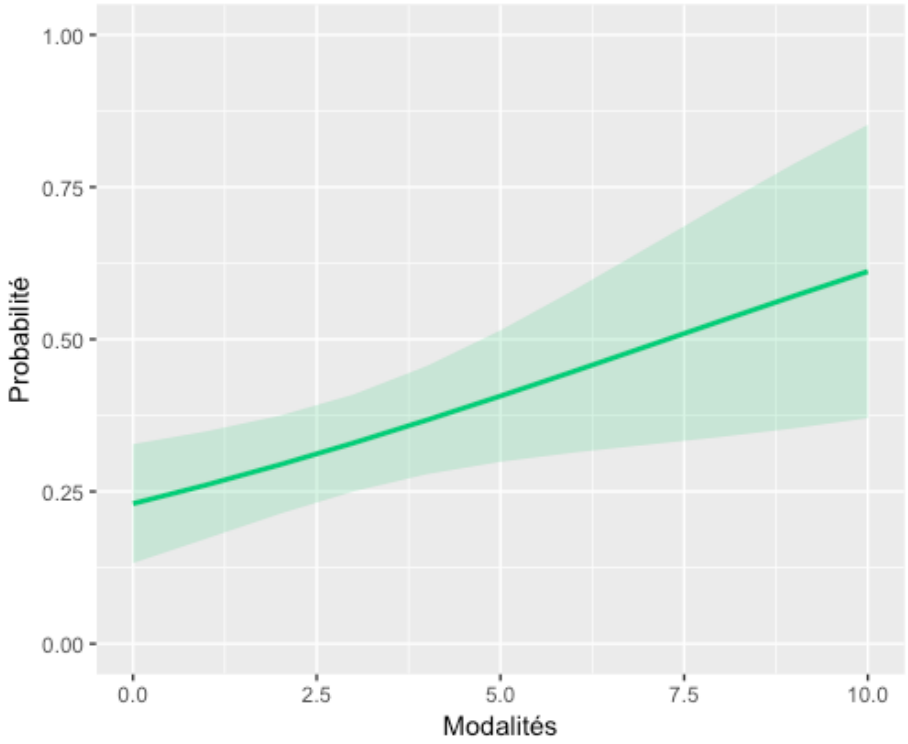


Figure 55. Régression logistique : probabilité d'adoption selon le niveau d'Image de Soi

Tableau 49. Probabilité d'adoption en fonction du score d'Image de Soi

Modalité Image de Soi	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.23	.26	.29	.33	.37	.41	.45	.49	.53	.57	.61

Motivation Intrinsèque

β =.35, p < .001

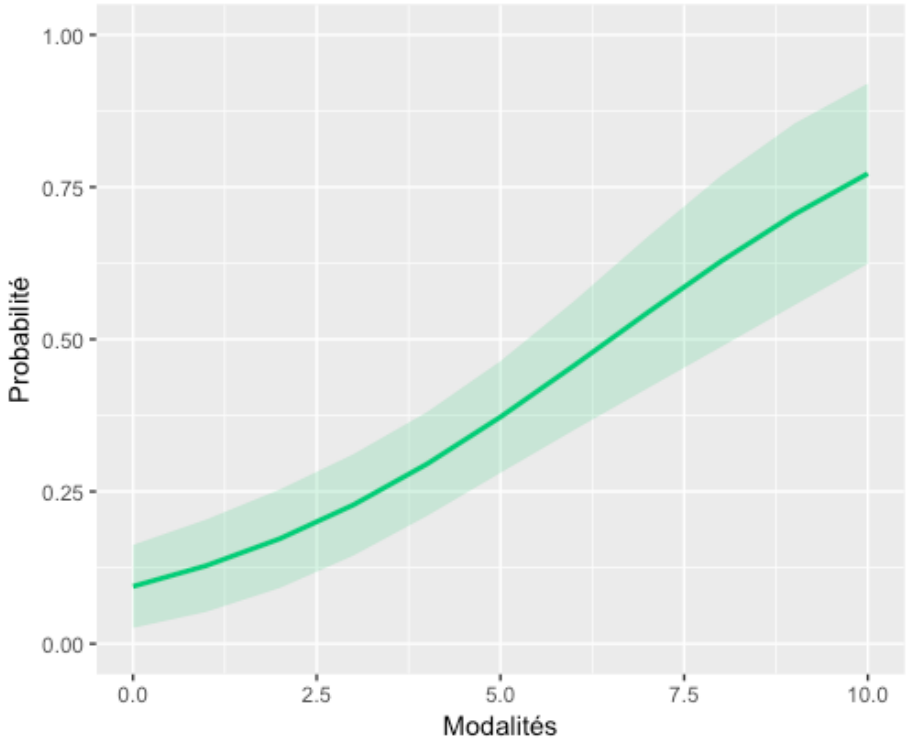


Figure 56. Régression logistique : probabilité d'adoption selon le niveau de Motivation Intrinsèque

Tableau 50. Probabilité d'adoption en fonction du score de Motivation Intrinsèque

Modalité Motivation Intrinsèque	0	1	2	3	4	5	6	7	8	9	10
Probabilité	.10	.13	.17	.23	.30	.37	.46	.54	.63	.71	.77

Annexe 19. Modélisation avant usage à l'aide de variables de second ordre

Au sein du Chapitre 1 : Modélisation de l'adoption avant usage, des modélisations de l'Intention d'Usage ont été réalisées. Cependant, les modélisations mises en place sont relativement simplistes avec uniquement des liens supposés directs entre les variables mesurées et l'Intention Comportementale. Une modélisation, plus complexe, a été envisagée et les résultats sont présentés ci-après.

1. Variables évaluées

Selon la littérature, un certain nombre de variables jouerait un rôle dans le jugement et l'adoption des produits technologiques :

- L'Utilité perçue et l'Utilisabilité perçue influenceraient l'Intention d'Usage (Dwivedi et al., 2011 ; King & He, 2006 ; Ma & Liu, 2004 ; Schepers & Wetzels, 2007) ainsi que le Jugement global (e.g., Aranyi & van Schaik, 2015a ; Hassenzahl, 2004 ; Mahlke, 2007 ; Mahlke et al., 2007 ; van Schaik & Ling, 2011)
- La Stimulation aurait une influence significative sur le Jugement global (Aranyi & van Schaik, 2015a ; van Schaik & Ling, 2008)
- La Confiance aurait un effet positif significatif sur l'Intention Comportementale (Corbitt et al., 2003 ; Gefen et al., 2003 ; Jarvenpaa et al., 2000 ; Pavlou, 2003)
- L'Identification serait un prédicteur du Jugement (van Schaik & Ling, 2008, 2011) et de l'Acceptabilité (J. D. Jackson et al., 2013 ; Kwon & Chon, 2009)
- La Motivation jouerait un rôle significatif sur l'Intention d'Usage (Koufaris, 2002 ; Lee et al., 2005)
- L'Utilisabilité perçue aurait un rôle positif sur la Motivation (Balog & Pribeanu, 2010 ; Lee et al., 2005)
- La Motivation Intrinsèque serait une variable médiatrice entre la Stimulation, l'Identification et l'Intention d'Usage (van Schaik & Ling, 2011)
- L'Influence Sociale serait un facteur important dans l'Adoption des produits technologiques (Dwivedi et al., 2011 ; Schepers & Wetzels, 2007 ; Taiwo & Downe, 2013)

- Effet de l'Influence Sociale sur la Confiance a été remarqué (Seckler, Heinz, Forde, Tuch, & Opwis, 2015)

L'acceptabilité et l'UX sont actuellement deux champs théoriques distincts. Peu de chercheurs se sont intéressés conjointement aux relations entre les variables de ces deux domaines. Van Schaik et ses collaborateurs (Aranyi & van Schaik, 2015a ; van Schaik & Ling, 2011) ont montré une influence importante de l'UX, notamment des qualités non fonctionnelles sur l'acceptabilité. Cependant, ces recherches restent rares et présentent certaines limites. Premièrement, elles ont majoritairement été conduites sur des produits comportant des qualités non-fonctionnelles relativement faibles : sites d'informations et Wikipedia (Aranyi & van Schaik, 2015a ; van Schaik & Ling, 2011). Ces systèmes sont en effet majoritairement centrés sur des qualités fonctionnelles. De plus, à notre connaissance, il n'existe pas d'études incluant conjointement les variables de l'UX et les facteurs affectivo-motivationnels présentés préalablement. Néanmoins, ces variables devraient influencer le comportement utilisateur et faire l'objet de nouvelles études (Saariluomaand & Jokinen, 2014). Ainsi, une approche intégrative pourrait offrir de nouvelles possibilités pour comprendre le jugement utilisateur (voir Figure 57).

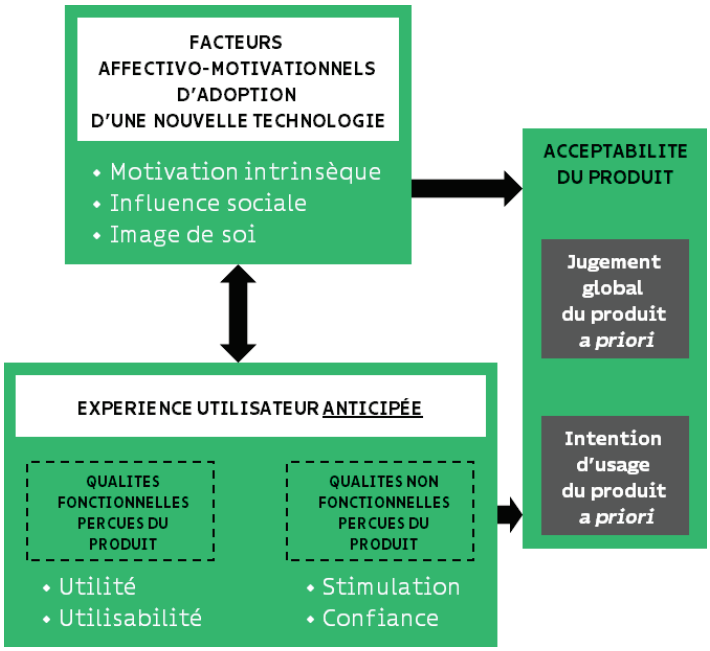


Figure 57. Modèle intégré de l'acceptabilité, de l'UX et des facteurs affectivo-motivationnels

La présente étude vise à prendre en compte l'ensemble de ces limites dans l'objectif d'investiguer l'influence des facteurs de l'UX et affectivo-motivationnels sur le jugement et l'acceptabilité. Pour réaliser cet objectif, des SEMs seront utilisées comme recommandé par Law et van Schaik (2010).

Selon les précédentes études, de nombreux liens entre les construits théoriques sont supposés. Ainsi, pour simplifier la modélisation statistique et par conséquent sa compréhension, des modèles avec des variables de second ordre sont estimés (F. F. Chen, Sousa, & West, 2005). Ce type de modèle a l'avantage d'être plus parcimonieux (Rindskopf & Rose, 1988) et ainsi d'offrir davantage de flexibilité dans la modélisation. Les modèles avec des facteurs d'ordre supérieur sont généralement utilisés lorsque les facteurs au premier ordre sont corrélés et peuvent considérés comme issus d'un même facteur d'ordre supérieur (F. F. Chen et al., 2005). Dans la présente étude, il semble possible de considérer les qualités des produits (Utilisabilité, Utilité, Stimulation et Confiance) comme issues d'un seul et même facteur de second ordre : l'UX.

Il pourrait être considéré de même avec les variables affectivo-motivationnelles (Motivation Intrinsèque, Image de Soi et Influence Sociale). Cependant, l'utilisation d'un facteur de second ordre requière au moins quatre items au premier niveau (F. F. Chen et al., 2005 ; Rindskopf & Rose, 1988). Par conséquent, les variables affectivo-motivationnelles ne peuvent être analysées qu'au premier niveau dans cette étude. Ainsi, le modèle suivant sera testé : voir Figure 58.

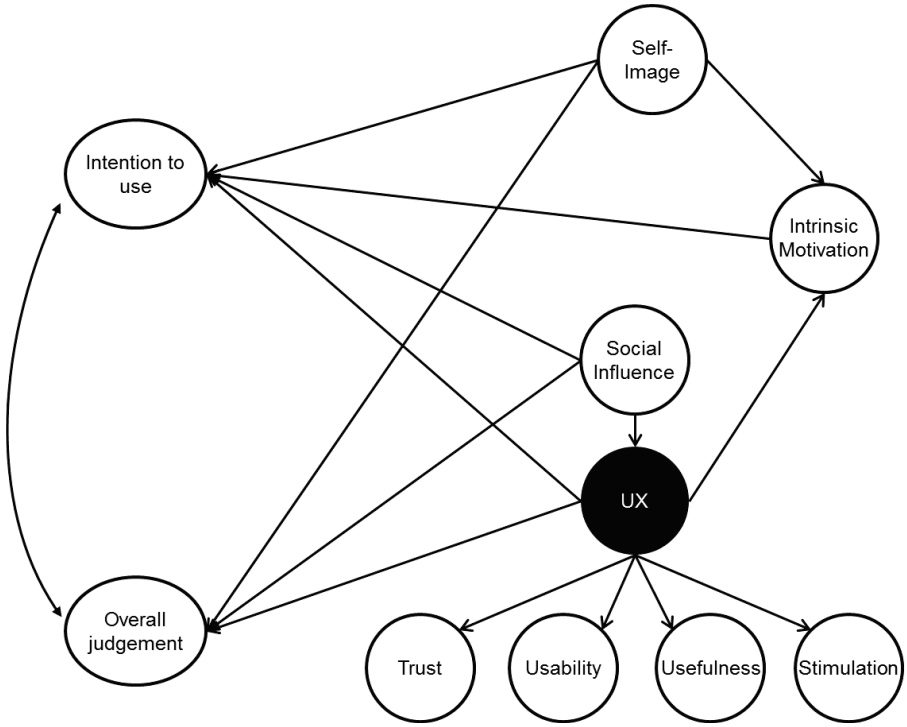


Figure 58. Modèle de recherche

2. Méthode et résultats

Les données utilisées pour cette modélisation sont identiques à celles présentées dans le Chapitre 1 : Modélisation de l'adoption avant usage, et par conséquent ne seront pas de nouveau présentées ici.

2.1. Analyses factorielles

Bien que les mêmes données que dans le Chapitre 1 : Modélisation de l'adoption avant usage, soient utilisées, il semble nécessaire de présenter les nouvelles analyses factorielles prenant en compte l'hypothèse d'un facteur de 2nd ordre pour l'échelle UX. Les analyses factorielles sur l'échelle A-M ne seront pas à nouveau présentées. Ainsi, pour l'échelle UX, les structures factorielles suivantes ont été comparées :

- M0 : un modèle à un facteur où l'unique facteur sature tous les items
- M1 : un modèle à deux facteurs corrélés correspondant aux qualités fonctionnelles (i.e., items rattachés à l'Utilité et l'Utilisabilité) et non fonctionnelles (i.e., items rattachés à la Stimulation et la Confiance).

- M2 : un modèle avec un facteur latent de 2nd ordre et 4 facteurs latents de 1er ordre correspondants aux facteurs définis a priori (i.e., Utilité, Utilisabilité, Stimulation et Confiance) et le facteur de 2nd ordre saturant les facteurs de premier ordre.
- M3 : un modèle à 4 facteurs corrélés correspondant aux facteurs définis a priori (i.e., Utilité, Utilisabilité, Stimulation et Confiance perçue)

La comparaison des modèles et l'ajustement des modèles indiquent que le modèle M2 comprenant un facteur de 2nd ordre s'ajuste le mieux aux données : voir Tableau 51.

Tableau 51. Ajustements et comparaisons des modèles pour les analyses factorielles de l'échelle UX

Modèle	#par	χ^2	df	$\frac{\chi^2}{df}$	CFI	RMSEA	SRMR	BIC	$\Delta\chi^2$	Δdf	p
M0	48	1532.28	104	14.73	.543	.209	.165	20387			
M1	49	1158.26	103	11.25	.663	.181	.164	19817	248.38	1	<.001
M2	52	137.35	100	1.37	.988	.034	.037	18278	233.70	3	<.001
M3	54	132.82	98	1.36	.989	.034	.030	18283	4.75	2	.093

2.2. Modélisation de l'Intention Comportementale

Après s'être assuré de la comptabilité de la structure factorielle avec les données, la modélisation de l'Intention Comportementale a été réalisée. L'estimation du modèle comprenant tous les liens supposés présente un ajustement satisfaisant (voir Tableau 51 – modèle M1). Cependant plusieurs liens apparaissent comme non significatifs. Afin d'obtenir le modèle le plus parcimonieux, de nouveaux modèles ont été estimés en supprimant un par un les liens non significatifs en s'assurant de la non dégradation de l'ajustement des modèles. Ainsi, les modèles suivants ont été comparés :

- M1 : le modèle incluant tous les liens supposés.
- M2 : suppression du lien entre Image de Soi et Motivation Intrinsèque.

- M3 : suppression du lien entre Image de Soi et Intention d'Usage.
- M4 : suppression du lien entre Influence Sociale et Intention d'Usage.
- M5 : fixation à 0 du lien entre Motivation Intrinsèque et Intention d'Usage[69].

Selon la comparaison des modèles et leurs ajustements respectifs (voir Tableau 52), le modèle M5 semble offrir le meilleur ajustement. Ce modèle (voir Figure 59) indique une influence significative de l'Image de Soi, de l'Influence Sociale et de l'UX sur le Jugement Global. En ce qui concerne l'Intention d'Usage, seul le facteur UX aurait une influence significative.

Tableau 52. Ajustements et comparaisons des modèles

Modèle	#par	χ^2	df	$\frac{\chi^2}{df}$	CFI	RMSEA	SRMR	BIC	$\Delta\chi^2$	Δdf	p
M1	112	696.76	448	1.56	.968	.042	.060	37599			
M2	111	698.45	449	1.56	.968	.042	.062	37596	1.82	1	.18
M3	110	699.21	450	1.55	.968	.042	.062	37591	.48	1	.49
M4	109	701.50	451	1.56	.967	.042	.062	37589	2.50	1	.11
M5[70]	111	704.59	452	1.56	.967	.042	.062	37585	2.88	1	.09

[69] Avec la suppression du lien entre l'Intention d'Usage et la Motivation Intrinsèque, la Motivation devient une variable de sortie (influencée par l'UX et n'influençant aucune variable). Par conséquent, plusieurs covariances sont estimées (i.e., avec le Jugement et l'Intention Comportementale – ajout de deux paramètres). Cependant, il n'y a pas d'hypothèse de covariance entre ces variables, ainsi pour éviter ce problème, le lien entre la Motivation Intrinsèque et l'Intention d'Usage est fixé à 0 plutôt que supprimé.

[70] Les effets directs et totaux pour l'Influence Sociale sont les suivants :

- Influence Sociale sur l'Intention d'Usage via la variable UX : β = .64 ; p < .001 ; *95% C.I* [.66 ; .89]
- Influence Sociale sur l'Intention d'Usage via la variable UX : β = .59 ; p < .001 ; *95% C.I* [.36 ; .57]
- Effet total de l'Influence Sociale sur le Jugement global : β = .44 ; p < .001 ; *95% C.I* [.24 ; .45]. L'effet total n'est pas calculé pour l'Intention d'Usage car il n'y a pas de lien direct.

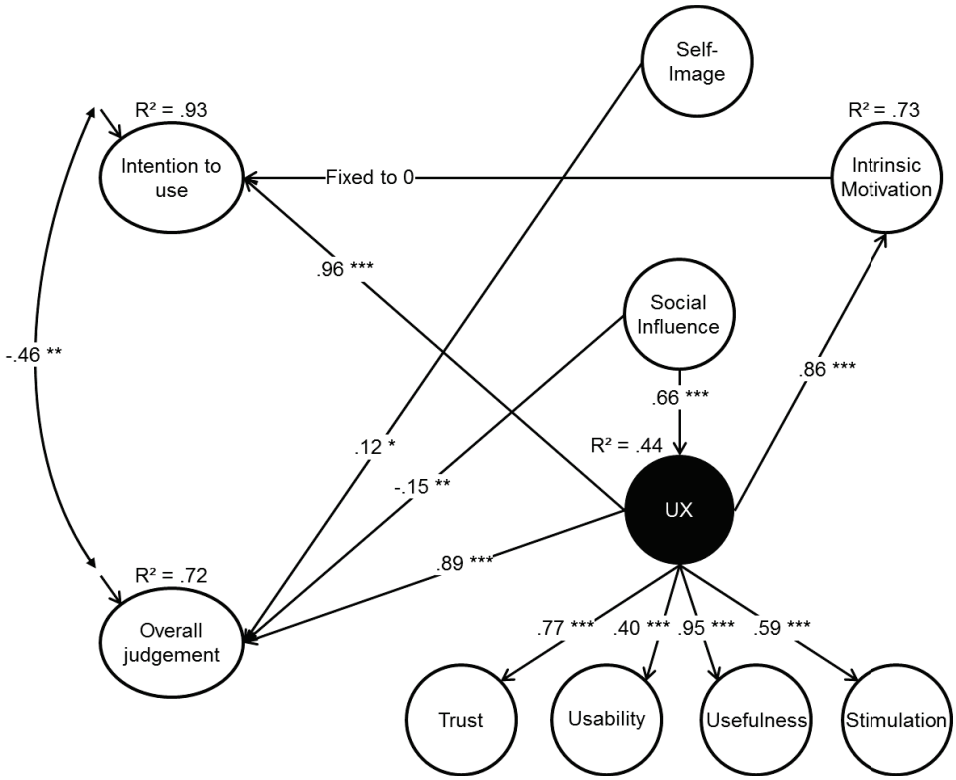


Figure 59. Modèle M5 avec les liens estimés[71]

3. Discussion et Conclusion

La structure factorielle supposée avec une variable de 2nd ordre pour l'UX est confirmée au sein des données. Cela signifie que les qualités fonctionnelles et non fonctionnelles pourraient être résumées par un seul facteur latent UX de haut niveau. Ce facteur UX est construit à partir de la variance de l'Utilité, l'Utilisabilité, la Stimulation et la Confiance. Ces variables (regroupées au sein du facteur UX) auraient un effet positif et significatif sur l'Intention d'Usage et le Jugement global envers une application mobile. Ce résultat est cohérent avec les études précédentes dans le contexte du jugement après usage (e.g., Aranyi & van Schaik, 2015a ; Dwivedi et al.,

[71] NS = Non significatif ; * = $p < .05$; $p < .01$; ***: $p < .001$. Les β sont totalement standardisés. Les variables observables ainsi que les covariances entre les variables ne sont pas représentées.

2011 ; King & He, 2006 ; Ma & Liu, 2004 ; Mahlke et al., 2007 ; Schepers & Wetzels, 2007).

Concernant les facteurs affective-motivationnels, il était supposé un effet de l'Image de Soi, de l'Influence Sociale et de la Motivation Intrinsèque. En ce qui concerne l'acceptabilité, seule l'Influence Sociale aurait un effet sur l'Intention Comportementale mais uniquement indirect. En effet, le lien entre l'Influence Sociale et Intention Comportementale serait totalement médiatisé par l'UX. Cet effet médiatisé de l'Influence Sociale sur l'Intention d'Usage a été démontré à plusieurs reprises dans le contexte du jugement après usage (Schepers & Wetzels, 2007). Au sein de notre étude, l'UX est mise à l'échelle sur l'Utilité perçue, par conséquent la relation de médiation est cohérente.

L'absence d'effet direct de la Motivation Intrinsèque et de l'Image de Soi sur l'Intention Comportementale peut s'expliquer de plusieurs manières. Tout d'abord, King et He (2006) argumentent que le choix d'utiliser ou non un produit est majoritairement lié à la perception d'utilité. Dans notre étude, le facteur UX est mis à l'échelle sur la variable Utilité perçue. De plus, les variables Intention Comportementale et Utilité perçue sont très corrélées. Ainsi, cette relation pourrait absorber l'ensemble de la variance de l'Intention d'Usage. Deuxièmement, même si les items de motivation intrinsèque ont été construits spécialement pour l'évaluation avant usage, les items sont orientés sur le plaisir attendu dans l'interaction avec l'application (voir Annexe 4). Comme les participants n'ont jamais testé l'application, il est possible que l'évaluation de cette dimension soit difficile et par conséquent non prise en compte dans leur jugement. Enfin, l'absence d'effet de l'Image de Soi peut s'expliquer par le contexte volontaire dans lequel sont les participants. En effet, Venkatesh et ses collaborateurs (2003) ont montré un effet après usage uniquement dans un contexte d'usage obligatoire. D'autres études ayant montré une influence après usage ont été réalisées dans un contexte d'usage obligatoire (e.g., J. D. Jackson et al., 2013 ; Kwon & Chon, 2009).

Selon les modélisations réalisées, seule l'Influence Sociale aurait un effet (uniquement indirect) sur l'acceptabilité en dehors du facteur UX. Ainsi, introduire des facteurs affectivo-motivationnels ne semble pas augmenter la précision de la prédiction de l'Intention Comportementale avant usage. En effet, le facteur UX influence majoritairement l'acceptabilité.

Au sein de ces modélisations, la relation entre l'Intention d'Usage et le Jugement global a également été explorée. Au sein de la littérature, peu d'études sont, à notre connaissance, disponibles sur ce lien. Cependant, il est facilement concevable que le jugement global envers un système ne soit pas forcément corrélé à l'adoption de ce système. Même si ces variables sont reliées l'une à l'autre, la corrélation peut n'être que partielle. Par exemple, une personne peut avoir un bon jugement envers un produit sans pour autant l'adopter. Van Schaik et Ling (2011) ont mis en avant que le jugement de qualité (*goodness*) n'influence pas significativement l'intention d'utiliser un produit. Au sein de notre étude, les variables influençant l'intention d'usage sont différentes de celles influençant le jugement envers un produit. En effet, l'Intention d'Usage est seulement expliquée par les qualités perçues du produit (UX). En ce qui concerne le Jugement global, il est influencé par l'UX, la Motivation Intrinsèque, l'Influence Sociale et l'Image de Soi. Dans cette étude, les participants ne sont pas dans une situation naturelle, en effet, le produit qu'ils doivent évaluer est imposé, par conséquent des différences pourraient apparaître par rapport à un contexte naturel.

En conclusion, même si les modélisations réalisées montrent des résultats intéressants, elles présentent l'inconvénient d'être complexes dans leur mise en œuvre et par conséquent difficilement applicables. Afin de répondre aux contraintes du projet dans lequel s'inscrit cette thèse, seules des modélisations les plus simples seront utilisées.

Index des tableaux

Tableau 1. Influence de l'Utilité et de l'Utilisabilité sur l'Intention d'Usage selon 4 méta-analyses 18
Tableau 2. Pourcentages de relations significatives à travers différentes études issues d'une méta-analyse sur l'UTAUT (Dwivedi et al., 2011) 22
Tableau 3. Statistiques descriptives de l'échelle UX sur l'échantillon Sleep Better .. 81
Tableau 4. Statistiques descriptives de l'échelle A-M sur l'échantillon Sleep Better. 81
Tableau 5. Statistiques descriptives Intention d'usage sur l'échantillon Sleep Better 82
Tableau 6. Comparaison des modèles d'AFC pour l'échelle UX 84
Tableau 7. Comparaison des modèles d'AFC pour l'échelle A-M 85
Tableau 8. Fiabilité et corrélations entre les variables latentes 86
Tableau 9. Statistiques descriptives de l'échelle UX sur l'échantillon Google Fit 90
Tableau 10. Statistiques descriptives de l'échelle A-M sur l'échantillon Google Fit .. 91
Tableau 11. Statistiques descriptives de l'Intention d'usage sur l'échantillon Google Fit 91
Tableau 12. Évaluation de l'invariance pour l'échelle UX 93
Tableau 13. Comparaison des modèles d'invariances factorielles pour l'échelle A-M 94
Tableau 14. Statistiques descriptives de l'échelle UX 108
Tableau 15. Statistiques descriptives de l'échelle A-M 108
Tableau 16. Statistiques descriptives de l'Intention d'Usage 108
Tableau 17. Comparaison des modèles d'AFC pour l'échelle UX 110
Tableau 18. Comparaison des modèles d'AFC pour l'échelle A-M 111
Tableau 19. Statistiques descriptives (moyenne et écart-type) et comparaison statistique entre l'évaluation avant et après usage 124
Tableau 20. Comparaisons entre les échelles avant usage UX à items multiples, à items uniques corrigés et à items uniques 137
Tableau 21. Comparaisons entre les échelles avant usage A-M à items multiples, à items uniques corrigés et à items uniques[51] 137
Tableau 22. Corrélations entre les échelles UX après usage à items multiples, à items uniques corrigés et à items uniques[53] 140

Tableau 23. Corrélations entre les échelles A-M après usage à items multiples, à items uniques corrigés et à items uniques 140
Tableau 24. Résultats de la régression logistique sur l'Intention Comportementale avant usage 149
Tableau 25. Probabilité d'adoption en fonction du score d'Utilité perçue..... 150
Tableau 26. Probabilité d'adoption en fonction du score d'Utilisabilité perçue..... 151
Tableau 27. Statistiques descriptives des items de l'échelle A-M..... 219
Tableau 28. Statistiques descriptives des items de l'échelle UX 219
Tableau 29. Statistiques descriptives des variables observables de l'acceptabilité 220
Tableau 30. Statistiques descriptives des items de l'échelle A-M..... 221
Tableau 31. Statistiques descriptives des items de l'échelle UX 221
Tableau 32. Statistiques descriptives des variables observables de l'acceptabilité 221
Tableau 33. Résultats de l'AFE pour l'échelle UX..... 222
Tableau 34. Résultats de l'AFE pour l'échelle A-M[61] 223
Tableau 35. Résultats AFE pour l'Intention d'Usage[61] 223
Tableau 36. Statistiques descriptives des items de l'échelle A-M..... 224
Tableau 37. Statistiques descriptives des items de l'échelle UX 225
Tableau 38. Statistiques descriptives des variables observables de l'acceptabilité 225
Tableau 39. Résultats des AFE pour l'échelle UX 227
Tableau 40. Résultats des AFE pour l'échelle A-M..... 228
Tableau 41. Covariances entre les variables latentes pour l'échelle A-M après usage 230
Tableau 42. Résultats de la régression linéaire multiple sur les données après usage 232
Tableau 43. Résultats de modélisation PLS-SEM sur les données après usage[66]. 233
Tableau 44. Probabilité d'adoption en fonction du score d'Utilité perçue..... 236
Tableau 45. Probabilité d'adoption en fonction du score d'Utilisabilité perçue..... 237
Tableau 46. Probabilité d'adoption en fonction du score de Stimulation perçue..... 238
Tableau 47. Probabilité d'adoption en fonction du score de Confiance perçue 239
Tableau 48. Probabilité d'adoption en fonction du score d'Influence Sociale 240
Tableau 49. Probabilité d'adoption en fonction du score d'Image de Soi 241
Tableau 50. Probabilité d'adoption en fonction du score de Motivation Intrinsèque 242
Tableau 51. Ajustements et comparaisons des modèles pour les analyses factorielles de l'échelle UX..... 247

Tableau 52. Ajustements et comparaisons des modèles..... 248

Index des figures

Figure 1. Théorie de l'Action Raisonnée adaptée de Davis, Bagozzi et Warshaw (1989) 16
Figure 2. Théorie du Comportement Planifié adaptée de Ajzen (1991) 17
Figure 3. TAM adapté de Davis (1989) 18
Figure 4. UTAUT (Venkatesh et al., 2003) 21
Figure 5. Coefficients bêta moyens des variables issus de la méta-analyse sur l'UTAUT (Dwivedi et al., 2011) 23
Figure 6. Echelle de valence extraites du SAM (Bradley & Lang, 1994) 30
Figure 7. Modèle d'O'Brien & Toms (2008) 34
Figure 8. Modèle de l'engagement selon O'Brien & Toms (2010) 35
Figure 9. Modèle de l'UX selon Hassenzahl et Monk (2010) 44
Figure 10. Modèle de l'UX selon Mahlke & Thüring (2007) : le CUE-Model 46
Figure 11. Éléments influençant l'UX au cours de l'interaction selon Roto (2007) 48
Figure 12. Éléments de l'UX selon Arhippainen et Tähti (2003) 49
Figure 13. Modèle de Robert & Lesage (2011a) 50
Figure 14. QoS et QoE pour les DIMEs (W. Wu et al., 2009) 51
Figure 15. Modèle intégré de l'UX et de l'acceptabilité (van Schaik & Ling, 2011) 57
Figure 16. Modèle utilisé par Aranyi & van Schaik (2015a) pour l'étude de sites d'informations 58
Figure 17. Modèle 1 testé par Février (2011) 59
Figure 18. Modèle 2 testé par Février (2011) 60
Figure 19. Modèle proposé par Karapanos & al. (2009) 62
Figure 20. Les composantes du jugement utilisateur envers un produit technologique : Antécédents, Perceptions et Conséquences 65
Figure 21. Modèle intégratif et dynamique de l'acceptabilité/acceptation d'une nouvelle technologie 67
Figure 22. Captures d'écran de l'application Sleep Better 78
Figure 23. Résultats AFC pour l'échelle UX 84
Figure 24. Résultats AFC pour l'échelle A-M 86
Figure 25. Modèle avec les liens estimés (modèle le plus parcimonieux) dans l'échantillon Sleep Better 88

Figure 26. Captures d'écran de l'application Google Fit 89
Figure 27. Modèle avec les liens estimés (modèle le plus parcimonieux) dans les deux échantillons..... 95
Figure 28. Capture d'écran de l'application Waze..... 105
Figure 29. Histogramme des données de la variable intention comportementale... 107
Figure 30. Résultats de l'AFC pour l'échelle UX 110
Figure 31. Modèle avec les liens estimés (modèle le plus parcimonieux) 113
Figure 32. Captures d'écran de l'application Sleep Better 122
Figure 33. Résultats de la régression multiple mettant en œuvre notre proposition de modèle EDT' 125
Figure 34. Score vrai, manifestations observables et erreur de mesure..... 130
Figure 35. Modèle de mesure avec une variable latente et 3 items..... 130
Figure 36. Modèle de mesure avec une variable latente et un item : modèle non identifié et non estimable..... 132
Figure 37. Variable latente avec un seul item sans prise en compte de l'erreur de mesure : modèle identifié et estimable 133
Figure 38. Résultats de l'étude de Bangor, Kortum et Miller (2009, p. 119) : Adjectif choisi en fonction du score au questionnaire SUS 145
Figure 39. Captures d'écran de l'application Google Fit 147
Figure 40. Régression logistique : probabilité d'adoption selon le niveau d'Utilité perçue..... 150
Figure 41. Régression logistique : probabilité d'adoption selon le niveau d'Utilisabilité perçue..... 151
Figure 42. Concept d'estimation de la probabilité d'adoption à partir de la régression logistique et d'une nouvelle valeur..... 152
Figure 43. Modèle intégratif et dynamique de l'acceptabilité/acceptation d'une nouvelle technologie mis à jour au regard des résultats des études..... 158
Figure 44. Scree Plot sur les données récoltées avec l'échelle UX après usage ... 226
Figure 45. Scree Plot sur les données récoltées avec l'échelle A-M après usage.. 226
Figure 46. Résultats analyses factorielles confirmatoires pour l'échelle A-M après usage..... 229
Figure 47. Résultats de l'analyse factorielle *bootstrapped* pour l'échelle UX..... 234
Figure 48. Résultats de l'analyse factorielle *bootstrapped* pour l'échelle A-M[68] 234

Figure 49. Résultats de l'analyse factorielle *bootstrapped* pour l'Intention d'Usage[68] 235
Figure 50. Régression logistique : probabilité d'adoption selon le niveau d'Utilité perçue 236
Figure 51. Régression logistique : probabilité d'adoption selon le niveau d'Utilisabilité perçue..... 237
Figure 52. Régression logistique : probabilité d'adoption selon le niveau de Stimulation perçue..... 238
Figure 53. Régression logistique : probabilité d'adoption selon le niveau de Confiance perçue..... 239
Figure 54. Régression logistique : probabilité d'adoption selon le niveau d'Influence Sociale..... 240
Figure 55. Régression logistique : probabilité d'adoption selon le niveau d'Image de Soi 241
Figure 56. Régression logistique : probabilité d'adoption selon le niveau de Motivation Intrinsèque..... 242
Figure 57. Modèle intégré de l'acceptabilité, de l'UX et des facteurs affectivo-motivationnels..... 244
Figure 58. Modèle de recherche..... 246
Figure 59. Modèle M5 avec les liens estimés 249

ACCEPTABILITE, ACCEPTATION ET EXPERIENCE UTILISATEUR : EVALUATION ET MODELISATION DES FACTEURS D'ADOPTION DES PRODUITS TECHNOLOGIQUES

Résumé

La compréhension des dimensions centrales dans l'évaluation et l'adoption d'un produit technologique par les utilisateurs est un domaine de recherche en pleine expansion (Yeh & Teng, 2012). Les principaux précédents travaux ont été construits autour trois théories : l'acceptabilité/l'acceptation (i.e., variables à l'origine de l'adoption ; Dubois & Bobillier-Chaumon, 2010), l'expérience utilisateur (i.e., qualités fonctionnelles et non fonctionnelles influençant l'expérience avec un produit ; Hassenzahl, 2004) et les variables affectivo-motivationnelles (i.e., dimensions rattachées au fonctionnement humain influençant le jugement et le comportement). Cependant, à notre connaissance, aucune approche unifiée n'a été proposée. Ainsi, l'objectif de cette thèse était de créer un nouvel outil de mesure en mettant en lien ces travaux théoriques. Pour ce faire, plusieurs études ont été conduites aux diverses temporalités du jugement utilisateur. Les résultats ont, tout d'abord, mis en avant des qualités psychométriques satisfaisantes pour l'outil de mesure proposé. De plus, les modélisations statistiques ont montré un effet majeur de l'utilité perçue à travers les temps d'évaluation. Les autres variables empruntées à ces divers champs théoriques ont des effets différents en fonction de la temporalité du jugement. Enfin, des travaux ont été conduits pour offrir des échelles à items uniques, limitant les problèmes de modélisation classique des échelles courtes, afin de faciliter les futures expérimentations.

Mots clés

Outils de mesure, Acceptabilité, Expérience Utilisateur, Modèles à équations structurelles

ACCEPTABILITY, ACCEPTANCE AND USER EXPERIENCE: EVALUATION AND MODELING OF ADOPTION FACTORS OF TECHNOLOGICAL PRODUCTS

Abstract

The understanding of main dimensions influencing the evaluation and adoption of technological product by users is a growing up research field (Yeh & Teng, 2012). Majority of previous research was based on three theories: Acceptability/Acceptance (i.e., key variables in product adoption; Dubois & Bobillier-Chaumon, 2010), user experience (i.e., functional and non-functional qualities influencing experience with a product; Hassenzahl, 2004) and affective-motivational factors (i.e., dimensions related to human functioning influencing judgements and behaviors). Nevertheless, to our knowledge, no unified approach was proposed. Thus, the objective of the thesis was to create a new measurement tool linking these theoretical fields. For the purpose, several studies were conducted at several temporalities of user judgment. The results indicated satisfactory psychometrics proprieties for the proposed measurement tool. Moreover, the statistical modeling showed a main effect of perceived usefulness through evaluation time. Lastly, work has been done in order to propose single-items scales, limiting the classical modeling problems with short scales, to facilitate futures experimentations.

Keywords

Measurement tools, Acceptability, User eXperience, Structural Equation Modeling

UNIVERSITÉ RENNES 2

Laboratoire de psychologie cognition comportement communication EA 1285 LP3C

b com