HEC MONTRÉAL

Impact d'une expérience psychophysiologique négative sur la satisfaction des consommateurs dans un contexte d'épicerie en ligne

par

Caroline Giroux-Huppé

Sciences de la gestion

(Option Marketing)

Sous la codirection de

Sylvain Sénécal, Ph. D. et Pierre-Majorique Léger, Ph. D.

Mémoire présenté en vue de l'obtention

du grade de maîtrise ès sciences en gestion

(M. Sc.)

Avril 2019

Résumé

Un consommateur se remémorant ses expériences passées est sujet à des biais cognitifs, l’empêchant de se rappeler précisément l'occurrence de moments critiques (points de friction). Dans une interaction longue et complexe comme l’épicerie en ligne, identifier précisément ces points de friction est difficile pour le consommateur et n’a jamais été fait de façon précise et fiable à l’aide de mesures psychophysiologiques. Comme l’expérience vécue par le consommateur, en ligne et hors ligne, affecte sa satisfaction et ses intentions, il est important de comprendre sa satisfaction du processus entier, des attentes initiales à la satisfaction post achat.

Ce mémoire étudie l’impact de moments critiques négatifs dans l’expérience client sur la satisfaction dans un contexte d’épicerie en ligne. Il présente et teste aussi la nouvelle méthode de visualisation développée pour mesurer ces moments critiques, qui combine quantitativement la valence émotionnelle et l’intensité émotionnelle, afin d’identifier des points de friction dans le parcours client. Une expérience en laboratoire avec 45 participants, basée sur l’«expectation-confirmation theory» (Oliver, 1980), a permis de comparer les résultats obtenus à l’aide des mesures implicites (psychophysiologiques) et explicites (rapportées par le participant).

Les résultats contribuent à la littérature existante en proposant une nouvelle méthode implicite et en montrant le rôle médiateur de la satisfaction et l’importance des émotions négatives. Du côté pratique, cette étude propose un nouvel outil pour le UX, permettant aux professionnels d’identifier efficacement des points de friction implicites dans le parcours client. Elle donne également aux épiciers une meilleure compréhension de la satisfaction, perçue et vécue.

Mots clés : Satisfaction, Points de friction, Consommateur, Expérience Client, Épicerie en ligne, E-commerce, Expérience émotionnelle négative, Données psychophysiologiques implicites, Intentions, Médiation

Table des matières

Liste des tableaux et des figures

Liste des figures

Premier article

Deuxième article

Liste des tableaux

Liste des abréviations

IS : « *Information Systems* »

UX : Expérience utilisateur

UX : « *User Experience* »

Avant-propos

Afin d'avoir l'autorisation de rédiger ce mémoire par articles, l'approbation de la direction administrative du programme de M.Sc. de HEC Montréal a été obtenu. Ce mémoire a donc été rédigé sous la forme de trois articles complémentaires. Le premier article a été soumis et accepté à la conférence *HCI International 2019*, qui se déroulera à Orlando (Floride), en juillet 2019. Le second article sera intégré dans un livre en développement portant sur le UX (expérience utilisateur). Il explique de façon détaillée la méthode utilisée pour calculer et visualiser les points de friction dans le parcours client. Ce second chapitre complète bien le premier chapitre en expliquant pas à pas les étapes pour reproduire la méthode utilisée dans ce chapitre. Finalement, le troisième article sera soumis à *l'International Journal of Electronic Commerce*. L'accord des co-auteurs de tous les articles a été obtenu afin que ceux-ci puissent être présentés dans ce mémoire. Ce projet de recherche a été approuvé par le comité d'éthique en recherche (CER) de HEC Montréal le 12 mars 2018.

Remerciements

La rédaction de ce mémoire s'est avérée beaucoup plus difficile qu'envisagé au départ. Je me considère extrêmement chanceuse d'avoir pu effectuer mon mémoire au Tech3Lab et d'avoir pu bénéficier de l'encadrement et du support. Mon cheminement n'aurait pas été le même sans l'apport considérable de plusieurs personnes, qui ont été présentes du début à la fin.

Je tiens tout d'abord à remercier mes codirecteurs, Sylvain Sénécal et Pierre-Majorique Léger pour leur support constant. Un énorme merci pour votre patience, vos conseils, vos relectures et la confiance que vous m'avez témoignée. Être sous votre direction m'aura permis de grandir professionnellement et je me considère extrêmement chanceuse d'avoir des directeurs qui valorisent autant le travail de leurs étudiants.

Je souhaite également remercier l'équipe du Tech3Lab, donc Shang Lin, Carl St-Pierre, Bertrand, David ainsi que toutes les assistantes de recherche qui m'ont aidé durant ma collecte : Ariel, Salima, Julianne, Diana et Myriam. Un merci tout particulier à Emma, pour ton efficacité et ton travail exceptionnel, mais surtout ta bonne humeur.

De nature sociable, je n'aurais jamais réussi ce long processus sans avoir des amis à qui parler. Merci à Ariane, Émilie, Aurélie, Gabrielle et Mario de m'avoir supporté et écouté tout au long du processus. Merci également à Michael, Félix, Karl-David, Christophe, Robin, Mario, Éric, Yasmine et Florian d'avoir été les meilleurs collègues.

Je tiens également à remercier ma famille, tout particulièrement ma mère, Francine. Je remercie aussi le Conseil de recherche en sciences naturelles et génie (CRSNG) et Prompt pour avoir contribué au financement de cette étude.

Introduction

Problématique générale de l'étude

L'industrie du commerce en ligne est en essor constant et ce, à l'échelle mondiale. Les ventes en ligne, de plus de 2 milliards de dollars en 2017, sont prévues de doubler d'ici 2021 (McNair, 2018). Les ventes d'épicerie en ligne augmentent, elles aussi, mondialement, avec un taux de croissance annuel composé estimé à 17% entre 2017 et 2022 (Forrester, 2018). Toutefois, alors que 58% des consommateurs mondiaux ont effectué un achat en ligne pour des produits de mode, seulement 21% des consommateurs auraient acheté des produits frais (Nielsen, 2017).

Les consommateurs semblent réticents à adopter l'épicerie en ligne et préfèrent les épiceries traditionnelles. Les recherches antérieures évoquent plusieurs raisons, dont le besoin de toucher les aliments avant de les acheter (Citrin et al., 2003) et la réticence au risque (Huang & Oppewal, 2006). Toutefois, plus de 30% des consommateurs américains ont essayé l'épicerie en ligne au moins une fois (Statista, 2017) et lorsque l'on considère spécifiquement les milléniaux, cette proportion monte à 43% (The Hartman Group, 2017). Par conséquent, les épiceries traditionnelles doivent maintenant s'adapter s'ils ne veulent pas perdre leurs parts de marché dans cette industrie hautement compétitive. En effet, certains détaillants opérant uniquement en ligne, comme Amazon, détiennent déjà de grandes parts du marché (Kahn, 2018).

Comprendre la satisfaction ainsi que les sources d'insatisfaction n'est pas facile pour les épiciers, particulièrement dans le contexte de l'épicerie en ligne. Comparé à une transaction simple comme acheter un vêtement, faire son épicerie en ligne est une tâche beaucoup plus complexe, puisqu'elle requiert de multiples décisions et que le processus de commande implique parfois des calculs arithmétiques (Desrochers et al., à paraître). De plus, le temps passé sur un site d'épicerie en ligne est beaucoup plus long comparativement au temps moyen passé sur un site de commerce de détail. À titre d'exemple, aux États-Unis, en 2018, les consommateurs passent en moyenne six minutes par visite sur un site web de commerce de détail, lorsqu'ils sont à l'ordinateur (Salesforces

Research, 2019). Finalement, en plus d'être long et complexe, le processus est également omnicanal, puisque les consommateurs sont amenés à faire leurs achats en ligne, puis à ramasser leur commande en épicerie, ou à la faire livrer. Cette croissance de l'omnicanal (Anesbury et al., 2016) ainsi que la prédominance du canal en ligne est un nouveau défi de taille pour les épiciers.

Sachant que les consommateurs loyaux peuvent être jusqu'à cinq fois plus profitables que les nouveaux consommateurs (Gupta & Kim, 2007), il devient crucial pour les épiceries d'offrir un service exemplaire et uniforme à la fois en ligne et en magasin pour fidéliser les consommateurs (Alexandris et al., 2008). Convertir les clients magasinant en épiceries traditionnelles en clients en ligne est une opportunité pour les épiciers de faire compétition aux grands joueurs de ce monde en augmentant leurs parts du marché ainsi que leurs profits (Lee, 2013).

Ainsi, alors que plus de 5 000 milliards de dollars sont en jeu mondialement dans le domaine alimentaire (Forrester, 2018), il est de l'intérêt à la fois du domaine théorique et managérial d'avoir une meilleure compréhension du rôle de la satisfaction dans l'ensemble du processus ainsi que des sources d'insatisfaction critiques.

Problématiques spécifiques de l'étude

Cette étude s'intéresse à deux problématiques spécifiques : une dans le domaine de l'expérience utilisateur et l'autre dans le domaine du marketing.

Du côté du UX, plusieurs études ont montré qu'un consommateur a de la difficulté à se rappeler ses expériences passées avec exactitude (Cockburn et al., 2017; Fredrickson and Kahneman, 1993, Eich and Schooler, 2000). Ainsi, il est peu probable que dans une expérience aussi longue que celle de l'épicerie en ligne, un consommateur puisse se rappeler exactement le moment précis de l'occurrence des moments les plus négatifs de son parcours. Comprendre les sources d'insatisfaction critiques dans le parcours client en ligne est capital afin d'améliorer l'expérience vécue par les consommateurs et leur satisfaction, afin de favoriser leur adhésion au processus d'épicerie en ligne.

Les études antérieures démontrent que de multiples raisons peuvent influencer la mémoire des consommateurs. Plusieurs études effectuées par Kahneman suggèrent que la mémoire humaine est influencée par des moments « peaks », qui deviennent ensuite les seuls éléments mémorisés de l'expérience précédemment vécue (Kahneman et al., 1993; Kahneman et al., 1997; Redelmeier & Kahneman, 1996). Par exemple, la «peak-end rule» stipule que les derniers moments d'une expérience sont plus facilement remémorés lors qu'une évaluation rétrospective de celle-ci (Cockburn et al., 2015). Le « peak effect », lui, stipule que le consommateur a tendance à mieux se souvenir du moment le plus intense de l'expérience (Cockburn et al., 2015). Ces résultats concordent avec ceux d'autres études, suggérant que les émotions les plus intenses et les plus négatives sont plus facilement mémorables (Ariely, 1998; Baumeister et al., 2001). D'un autre côté, une étude récente suggère que les premiers et derniers moments ont un impact sur l'auto-évaluation des consommateurs de leurs propres émotions (Lourties et al., 2018), c'est-à-dire que les consommateurs se rappelleront plus facilement du début et de la fin de leur expérience que du milieu. Ainsi, l'effet de primauté viendrait également affecter la mémoire du consommateur (Lourties et al., 2018). Par conséquent, il est tout à fait normal que lorsque l'on demande à un consommateur de se rappeler d'un moment précis d'une expérience passé, celui-ci puisse avoir peu confiance en sa réponse (Kahneman et al., 1997).

Les mesures explicites d'évaluation de l'expérience utilisateur (e.g. entrevues, questionnaires), basées sur une évaluation rétrospective humaine, peuvent être sujettes à des biais cognitifs (Cockburn, 2017; Cockburn 2015; Eich, 2000). Par conséquent, puisque les utiliser seules pourrait produire des résultats biaisés, il est hautement pertinent d'utiliser additionnellement des mesures implicites afin d'augmenter la validité des résultats obtenus (Ortiz de Guinea et al., 2013).

Du côté du marketing, il est primordial de comprendre l'impact réel de moments négatifs vécus lors d'une expérience client sur la satisfaction du consommateur et ses intentions, de réachat et de recommander, afin de juger de sa gravité. De plus, comprendre le rôle de la satisfaction du consommateur en ligne dans la relation entre les émotions vécues et les intentions est d'une grande valeur ajoutée pour les épiciers comme pour le milieu académique.

Le concept de satisfaction a été maintes fois utilisé dans la littérature en marketing, particulièrement dans le but de prédire le comportement des consommateurs (Oliver, 1993, Kim, Ferrin & Rao, 2009; Kim, 2012). Cependant, très peu d'études ont analysé le processus complet, des attentes avant l'achat à la satisfaction après avoir récupéré sa commande à l'épicerie. De plus, la plupart des modèles tentant d'expliquer la satisfaction utilisent des mesures explicites, c'est-à-dire basées sur l'expérience rétrospective perçue des consommateurs. Or, tel que démontré précédemment, les construits de mesures explicites s'appuyant sur la mémoire des consommateurs sont souvent biaisés. Par conséquent, intégrer des mesures psychophysiologiques pour mesurer les moments négatifs vécus lors d'une expérience client dans un modèle tentant d'expliquer la satisfaction permettrait de voir comment les mesures implicites et explicites interagissent entre elles dans le développement de la satisfaction finale (Ortiz de Guinea et al., 2014).

Ainsi, bien que provenant de domaines de recherche différents, ces problématiques sont complémentaires et génèrent plusieurs questions de recherche.

Questions de recherche

Alors que l'avènement de l'épicerie en ligne multicanale est encore récent, comprendre les raisons derrière l'aversion ou l'insatisfaction des consommateurs permettrait aux épiciers d'améliorer l'expérience vécue par les consommateurs et d'offrir une expérience unifiée et fluide. Il est donc crucial de comprendre les expériences négatives réellement vécues par les consommateurs afin de proposer les solutions appropriées.

Ce mémoire par articles explore l'impact d'intenses émotions psychophysiologiques négatives sur la satisfaction et les intentions des consommateurs dans un contexte d'épicerie en ligne. Dans un premier temps, le parcours des consommateurs est examiné quantitativement afin de déterminer les moments où l'expérience vécue par les consommateurs est à la fois intense et négative. Ces moments négatifs sont également comparés avec les perceptions des consommateurs suite à l'expérience afin de déterminer s'ils sont similaires ou non. L'intensité des émotions est mesurée à l'aide du niveau de sudation, capturé dans la paume de la main, alors que la valence émotionnelle, qui mesure le côté positif ou négatif des émotions, est mesurée à l'aide du logiciel d'analyse des émotions faciales FaceReader. Des données oculométriques sont également utilisées afin d'aller identifier dans l'enregistrement de l'expérience les moments précis où les consommateurs ont ressenti une émotion intense négative, afin de tenter d'en comprendre la cause.

Ce mémoire est divisé en deux parties. Dans la première partie, les articles 1 et 2 tentent de répondre aux questions de recherche suivantes :

Q1: Comment les mesures psychophysiologiques peuvent-elles nous permettre d'identifier les points de friction implicites dans le parcours client en ligne?

Q2: Les consommateurs sont-ils en mesure d'identifier rétrospectivement avec précision les points de friction de leur propre parcours en ligne?

Dans la deuxième partie, lors de la même étude, les émotions négatives implicites, identifiées grâce à la méthode précédemment décrite, sont combinées à des mesures explicites afin de tenter de répondre aux questions suivantes :

Q3 : Quel est le rôle de la satisfaction dans la relation entre l'expérience utilisateur du consommateur et ses intentions ?

Q4 : Quel est l'impact d'une expérience utilisateur négative sur la satisfaction du consommateur ?

Les résultats des deux dernières questions sont présentés dans l'article 3 du mémoire.

Objectifs de l'étude, contributions, et implications

Ce mémoire vise deux objectifs et est réalisé en deux temps. Le premier objectif est d'introduire et de tester une nouvelle méthode de mesure implicite de l'expérience d'un consommateur. Le second est de combiner les résultats de celle-ci avec des mesures explicites, afin de tenter de mieux expliquer la satisfaction dans un contexte d'épicerie en ligne. D'un point de vue théorique, ce mémoire à des contributions autant dans le domaine de l'expérience utilisateur que du marketing. Du point de vue de l'expérience utilisateur, celui-ci contribue à la fois à la théorie et à la pratique en proposant une méthode fiable permettant de visualiser les réactions émotionnelles les plus intenses vécus par les consommateurs lorsque ceux-ci performent une tâche. Cette méthode permet ainsi d'identifier des points de friction critiques dans le parcours du consommateur, qui sont plus précis et plus fidèles que ceux mentionnés par le consommateur après la tâche. Du point de vue du marketing, celui-ci propose un modèle de recherche basé sur la littérature existante, intégrant à la fois des mesures implicites et explicites, permettant une meilleure compréhension de l'interaction entre ces deux types de mesures lors de la formation de la satisfaction. Cette étude a ainsi démontré l'effet médiateur de la satisfaction en ligne perçue par les consommateurs, médiant l'effet entre l'expérience négative vécue et les intentions. D'un point de vue pratique, ces résultats permettront aux détaillants de mieux cibler les points de friction critiques de leurs consommateurs et d'avoir une meilleure compréhension des différents facteurs affectant la satisfaction de ceux-ci.

Informations sur les articles

La collecte de données en laboratoire de ce mémoire a été réalisée à l'hiver 2018 et à l'hiver 2019, alors que l'étudiante de ce mémoire travaillait sous la bourse de la Chaire de recherche industrielle CRSNG-Prompt en expérience utilisateur. Les résultats obtenus ont permis la rédaction de trois articles. Deux articles de ce mémoire ont été rédigés suite à l'analyse de la première collecte de données, où 21 participants ont été amenés en laboratoire. Un de ceux-ci, présenté au chapitre 2, sera présenté à la conférence *HCI International 2019.* (*Human-Computer Interaction International Conference*) à Orlando (Floride) en juillet 2019. L'autre, présenté au chapitre 3, sera intégré dans un livre en développement portant sur le UX (expérience utilisateur). Le troisième article, présenté au chapitre 4 dans ce mémoire, a été rédigé suite à une seconde collecte de données identique à la première, où 24 autres participants ont été collectés afin d'augmenter la taille de l'échantillon. Celui-ci est en préparation pour une soumission à l'*International Journal of Electronic Commerce*.

Par conséquent, ce mémoire par articles est présenté en deux parties. La première partie du mémoire, composée des chapitres 2 et 3, se concentre sur la nouvelle méthode d'évaluation de l'expérience utilisateur, développée dans le cadre de ce mémoire, qui permet de comparer l'expérience perçue des utilisateurs avec celle réellement vécue.

Alors que le chapitre 2 présente les résultats obtenus à l'aide de cette méthode, le chapitre 3 offre un complément d'information en permettant de reproduire pas à pas les étapes nécessaires pour utiliser la méthode. La seconde partie du mémoire, composée du chapitre 4, se concentre sur le rôle médiateur de la satisfaction et sur l'impact d'une expérience négative vécue, mesurée à l'aide de la méthode développée précédemment, sur la satisfaction du consommateur et ses intentions.

Ainsi, l'ordre des présentations des articles ainsi que les quelques redondances qui s'y retrouvent s'expliquent par le fait qu'il s'agisse d'un processus itératif. En effet, le premier article, soumis à HCI, a permis l'obtention de commentaires constructifs de la part des arbitres. Ces commentaires ont permis d'avoir la rétroaction nécessaire pour améliorer le troisième article, qui sera soumis à l'*International Journal of Electronic Commerce*.

Résumé du premier article

Un point de friction psychophysiologique se définit comme étant un instant où l'utilisateur est à la fois dans un état d'activation émotionnelle élevé et ressent une valence émotionnelle négative, par rapport à son état au repos. L'objectif du premier article était de proposer une nouvelle méthode de visualisation permettant d'identifier les points de friction implicites dans l'expérience du consommateur en ligne ainsi que de démontrer comment les mesures psychophysiologiques utilisées permettent d'identifier avec précision où se situent les points de friction dans le parcours de celui-ci. L'étude compare également les points de friction mentionnés par les consommateurs rétrospectivement avec ceux réellement vécus durant l'expérience afin de voir si les consommateurs sont capables d'identifier fidèlement les points de friction qu'ils ont eux-mêmes vécus. En identifiant ces points de friction implicites et en les combinant avec des données oculométriques, il est possible de comparer les points de friction communs aux consommateurs et de comprendre plus en profondeur les sources de points de friction, particulièrement ceux que les consommateurs n'ont pas réussi à identifier rétrospectivement. Cette expérience en laboratoire a été menée auprès de 21 participants, répartis en trois groupes égaux sur trois sites différents d'épicerie en ligne. Les résultats démontrent que les points de friction psychophysiologiques implicites peuvent être identifiés avec précision et sont plus fiables que ceux mentionnés explicitement par les consommateurs rétrospectivement. Les résultats démontrent également qu'une très faible proportion des points de friction sont identifiés par les consommateurs. Ces résultats contribuent à la littérature sur l'expérience utilisateur en proposant une méthode fiable permettant de visualiser les réactions émotionnelles des consommateurs durant une interaction avec un site web et en démontrant l'apport que peuvent avoir les points de friction implicites autant au niveau de la recherche que de la pratique.

Résumé du deuxième article

Le deuxième élément composant le mémoire est un chapitre de livre expliquant étape par étape la méthode à utiliser pour calculer, visualiser, identifier et comparer les points de friction psychophysiologiques implicites définis dans l'article précédent. La méthode utilisée permet d'identifier systématiquement les points de friction, de façon plus précise, autant au niveau du temps que du contenu. Pour illustrer la méthode proposée, le contexte de l'épicerie en ligne a été utilisé. Une expérience en laboratoire a été menée auprès de 21 participants, répartis en trois groupes égaux sur trois sites différents d'épicerie en ligne. Les bénéfices de la méthode comparée aux méthodes plus traditionnelles sont expliqués, ainsi que les outils et logiciels nécessaires pour la répliquer le plus fidèlement possible.

Résumé du troisième article

Alors que l'épicerie traditionnelle reste la préférée des consommateurs, l'épicerie en ligne est en pleine croissance et ce, globalement. Avec l'émergence de multiples canaux, le commerce en ligne se démarque, permettant ainsi aux consommateurs d'effectuer un parcours à la fois en ligne et hors ligne, en sélectionnant les produits désirés en ligne pour ensuite aller les chercher en magasin. Les consommateurs s'attendent donc à obtenir une expérience fluide avec le même niveau de service en ligne et hors ligne. Cependant, les épiciers, peu habitués à la place que prend le commerce en ligne dans ce secteur plutôt traditionnel, éprouvent présentement des difficultés à satisfaire leurs consommateurs. Cet article tente de combler un manque dans la littérature en étudiant le processus d'achat complet du consommateur, des attentes avant achat à la satisfaction suite à la collecte de la commande en magasin. De plus, dans la partie en laboratoire de l'étude, des mesures psychophysiologiques implicites sont utilisées afin de pallier au fait que l'évaluation rétrospective d'une expérience par le consommateur peut être sujette à de nombreux biais, tel que montré dans les articles précédents. Afin d'obtenir une meilleure compréhension du processus, une expérience multicanale a été menée auprès de 45 participants. Ceux-ci ont complété la première partie de l'expérience en laboratoire et la seconde en épicerie, permettant la collecte de mesures implicites et explicites. Les résultats démontrent un effet médiateur de la satisfaction entre l'expérience négative vécue par le consommateur et ses intentions, ainsi que l'impact significatif d'une expérience vécue négative sur la satisfaction et les intentions. Cette étude contribue à la fois à la littérature en marketing et en commerce électronique, en démontrant le rôle médiateur de la satisfaction et l'importance des émotions négatives lors de l'expérience vécue.

Afin de résumer la contribution personnelle de l'étudiante dans la rédaction de chacun des articles de ce mémoire effectué au Tech3Lab, un tableau récapitulatif (voir Tableau 1) est présenté. Les contributions pour chacune des parties sont exprimées en pourcentages.

Tableau 1 – Contributions dans la rédaction des articles

Étape du processus	Contribution
Définition des besoins du partenaire et des questions de recherche	Identifier les besoins d'affaire du partenaire et développer des questions de recherche en lien avec ceux-ci – 70% • Les besoins d'affaires ont été établis avec l'aide de Pierre-Majorique Léger, Sylvain Sénécal et l'équipe du Tech3Lab. • La réflexion derrière les questions de recherche et les construits à tester a été appuyée par mes directeurs de recherche.

Revue de la littérature	Rechercher et lire des articles scientifiques reliés aux différents construits testés – 100% Rédiger la revue de la littérature – 100% ● Relecture et correction des articles par mes directeurs de recherche Sélectionner les outils de mesures à utiliser en fonction des objectifs de recherche-80% ● L'équipe du laboratoire a validé les outils de mesures physiologiques et s'est assuré que ceux-ci étaient les plus cohérents en fonction des objectifs de recherche.
Conception du design expérimental	Effectuer la demande au CER – 100% ● L'équipe du laboratoire a appuyé ma démarche et s'est assuré que la demande soit complète et adéquate. ● Utilisation des formulaires de compensation et de consentement standardisés du Tech3Lab Concevoir le protocole d'expérimentation – 100% ● L'équipe du laboratoire a révisé le protocole pour approbation avant la collecte Organiser la salle de collecte – 90 % ● L'équipe du laboratoire a installé le matériel nécessaire. Mener des prétests afin de s'assurer de la fluidité de l'expérience et de la qualité des données – 100%

	● Les assistantes de recherche ont été présentes lors de tous les prétests afin de s'assurer de la qualité des données.
Recrutement des participants	Déterminer les conditions de participation et rédiger le questionnaire de recrutement – 100 % Recruter et gérer les participants au travers du panel – 100 % Administrer les compensations, pour les phases 1 et 2 – 100 % Faire un suivi avec les participants pour la phase 2 de l'expérience -90% ● Le cartable d'expérience a été assemblé par l'équipe du laboratoire.
Collecte de données	Collecter les données -50% ● Les données ont été collectées avec l'assistance des assistantes de recherche et ce, pour toutes les collectes.
Extraction et transformation des données	Extraire les données physiologiques (oculométrie, EDA, EKG) et reportées par questionnaires (Qualtrics) afin d'effectuer des analyses statistiques -100% Mettre en forme et regrouper les différents fichiers de données afin d'effectuer des analyses statistiques -100%
Analyse des données	Analyser des données oculométriques – 100 % Analyser des données physiologiques-100% Retranscrire et analyser les verbatims des entrevues-80%

	● Obtention de l'aide de l'équipe de laboratoire pour la retranscription des entrevues Analyser les statistiques du mémoire – 80% ● Interprétation des résultats statistiques afin d'en tirer des conclusions ● Obtention de l'aide d'un statisticien pour effectuer les tests statistiques plus complexes à l'aide du logiciel SAS 9.4 et SPSS
Rédaction	Écrire les articles présents dans le mémoire- 100% ● Mes directeurs de recherche ont amélioré significativement la qualité des articles à l'aide de leurs commentaires.

Structure du mémoire

Les trois prochains chapitres de ce mémoire présentent les résultats obtenus lors de cette étude, collectée à deux reprises, de façon identique. Par conséquent, il est à noter que le design expérimental, ainsi que la méthode restent les mêmes entre les articles. Seul le nombre de participants impliqués diffère. Finalement, le chapitre 5 de ce mémoire présentera une synthèse des résultats obtenus ainsi que les contributions aux domaines de l'expérience utilisateur et du marketing. Les limitations de l'étude ainsi que les futures avenues de recherche seront également abordées.

Chapitre 2: Premier Article

Identifying Psychophysiological Pain Points in the Online User Journey: The Case of Online Grocery[1]

Caroline Giroux-Huppé[1], Sylvain Sénécal[1], Marc Fredette[1], Shang Lin Chen[1], Bertrand Demolin[1] , and Pierre-Majorique Léger[1]

[1] HEC Montréal, Montréal, Canada
caroline.giroux-huppe@hec.ca

Abstract.

The objective of this study is to identify implicit psychophysiological pain points during an e-commerce interaction. In this article, we propose a method that allows to identify implicit pain points in the user's experience, by targeting moments when the user has both a high level of arousal and a negative emotional valence, compared to his baseline state; which means that the user feels an intense negative emotion. Identifying those pain points and combining them with eye-tracking data gives key insights into the user journey and helps identify implicit pain points shared among users. It also allows to gain a deeper understanding of pain points that users may fail to identify during the post-task interview. Our results show that the temporal occurrence of psychophysical pain points can be accurately identified and that it is more reliable than pain points explicitly mentioned by users. This study

[1] Giroux-Huppé, C.; Léger, P.-M..; Fredette, M.; Sénécal, S. « Identifying Psychophysiological Pain Points in the Online User Journey: The Case of Online Grocery.» In *International Conference on HCI in Business, Government, and Organizations* (accepted). Springer, Cham.

contributes to the user experience literature and practice by proposing a reliable method to visualise peak emotional reactions experienced by users while performing a task. Thus, providing more precision and reliability in identifying pain points when compared to pain points mentioned by users after the task.

Keywords: User Journey ·Visualisation Method · Psychophysiological Pain Points · Online Grocery Shopping

Introduction

User experience is a user's perceptions and responses resulting from the use of an interactive system, including emotions, beliefs, preferences, physiological responses and much more (ISO, 2010). To measure these responses, most UX research focuses on explicit methods such as questionnaires and interviews. For example, emotions, or user's feelings regarding a system, have previously been measured using a self-report scale developed by Hassenzahl et al (2003). However, it is difficult for users to precisely report on their own experience. Prior research shows that there is an important difference between what users felt during the experience and how they recalled it afterwards (Cockburn et al., 2017; Eich & Schooler, 2000). Recent findings suggest the influence of multiple biases, such as the peak effect, where the user tends to remember the most intense moment better, and the peak-end rule, where the user's impressions toward the experience tend to be influenced by the final moment (Cockburn et al., 2015). Furthermore, it has been shown that the intensity of emotions felt plays an important part in the recalling process (Ariely, 1998) and that negative memories tend to be better remembered than good ones (Baumeister et al., 2001).

Considering the lack of proper methods to accurately identify implicit pain points in the UX literature, we propose a systemic method that uses physiological data to identify pain points in a user online journey. Pain points can either be explicit, implicit, or both. An explicit pain point, usually derived from qualitative data, is defined as the negative emotion consciously felt by the participant during a particular moment in the task and mentioned by the participant during or after the task. It is commonly used in marketing research (Wang et al., 2016). An implicit pain point, however, is defined here as a moment, in reaction to an event during the interaction, during which the user experiences an automatic physiological activation characterized by a high level of emotional arousal and a negative emotional valence. Building upon previous research on peak loads, that identifies the exact moments users approach or pass their cognitive capacities (Mirhoseini et al., 2017), we use psychophysiological measures of emotional valence and arousal to build a metric that identifies pain points in the online user journey. We then illustrate the results using a journey map representation that allows a better understanding of the reasons behind those pain points as well as an easier comparison, either between different tasks or systems.

Gaining a deeper understanding of the reasons behind pain points contributes to HCI literature and practice by providing insights on peak emotional moments in users' experiences. It also allows UX designers to significantly improve their design by knowing precisely and accurately where the pain points are located, without interrupting users' authentic interactions with the website.

Literature Review

Current Methods to Assess Customer Experience

Customer experience contributes to the success of e-commerce websites and thus to a company's viability. Indeed, understanding customers and meeting their needs have been shown to be keys to success (Temkin, 2010). There is therefore a vast amount of literature focusing on analyzing the customer experience, using a variety of methods, such as personas, experience maps, blueprints, and walk-through audits (Johnston et al., 1990). However, these methods usually focus on a portion of customer experience, failing to give an overall picture. It has therefore been suggested that combining complementary methods offer a deeper understanding of user experience, while adding implicit measurement, such as physiological tools, allows for a more precise measure of the emotional journey of the participant (Nah & Xiao, 2018, p.327).

A first method, Customer Experience Modeling, has been developed in the service sector to better synthesize the whole customer journey and the sequence of the different touchpoints by using customer-centric soft goals (Teixeira et al., 2012). Soft goals are part of a goal-oriented analysis that allows problem detections in interactions by taking into account the subjective nature of the experience in the customer's evaluation of their different levels of satisfaction (Mylopoulos et al., 1999). It allows to discover pain points that emerge from interactions. Methods such as Customer Experience Modeling derive pain points from qualitative data. For example, the analysis of common words and sentences while completing a task (Wang et al., 2016).

Another method, the Customer Job Mapping, also known as the customer centered innovation map, consists of breaking down, step by step, every task customers face, in order to find new ways to innovate. Certain tasks or parts of tasks can bring difficulties for customers and are thus classified as pain points. The main difference between this method and Customer Experience Modelling is that this method focuses on what customers are trying to achieve at every step, instead of looking at what they are doing (Bettencourt & Ulwick, 2008).

The Customer Journey Map, a more recent method, is a diagram, illustrating every touchpoint a consumer has with the company, every step of the way and through every channel used across the company (Richardson, 2010). An example of a Customer Journey Map for grocery shopping can be found in Fig. 1. It is used both in the design service field, to help design the experience, and in the user experience field, to better understand the customer experience (Moon et al., 2016). Customer Journey Maps allow companies to focus on the entire customer experience rather than individual interactions (Rawson et al., 2013). However, recent research suggests that this method is still far from flawless, as it assumes all touchpoints are equally important to every customer, which is not the case (Rosenbaum et al., 2017). In order to identify the most important touchpoints, Customer Journey Maps should be linked with consumer research by using explicit measures such as self-administered questionnaires and interviews (Rosembaum et al., 2017). Another problem with Customer Journey Maps is that although they are now used in various industries, no clear process to design journey maps has been established, which makes it extremely difficult to compare across websites or interfaces, leading to inconsistent and non-generalizable results (Moon et al., 2016).

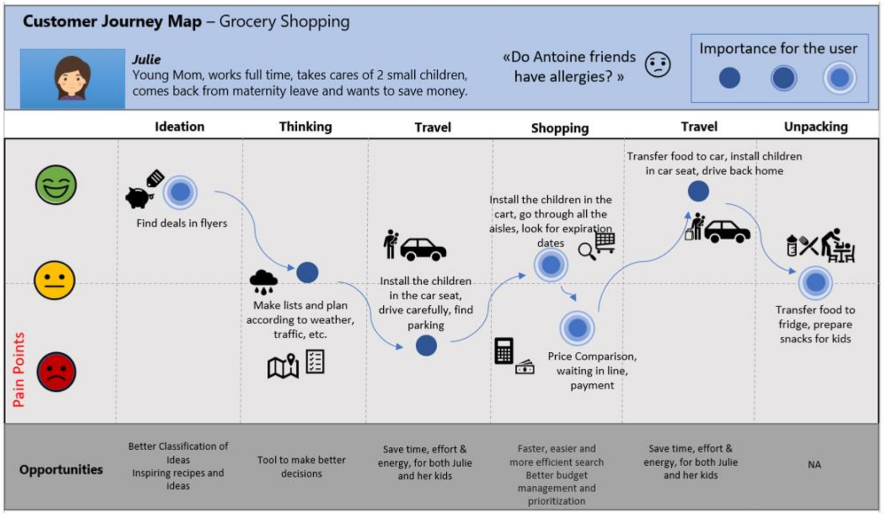


Fig. 1. Example of a Customer Journey Map

As seen from the above-mentioned methods, customer experience measurement has mainly been observed from a qualitative angle, using focus groups or observations, with the exception of surveys, which can include both qualitative and quantitative data (Moon et al., 2016; Nenonen et al., 20-8; Rawson et al., 2013). To understand the consumer's complete experience, data driven, quantitative analysis must be combined with qualitative, judgement driven evaluations (Rawson et al., 2013). There is currently a lack of methods that combine both these approaches (Rawson et al., 2013), as well as a lack of quantitative methods that would make experiences comparable across websites (Mangiaracina et al., 2009). To this day, there is no agreement on a method that would allow to evaluate all aspects of customer experience while reflecting reality, particularly when the user is completing complex tasks (Lallemand & Gronier, 2018). A recent study highlights the importance of using implicit measures to validate the data obtained from the participant's perceived emotions, to make sure all users emotions and reactions are considered (Alves et al., 2012). Another recent study used electroencephalography (EEG) and eye tracking to explore customer experience in order to develop new visualization methods (Alves et al., 2012). The authors quantified user experience with data such as attention levels, eye blinks, and pupil size. This study's difference compared to the previous ones is that the data collected comes from implicit, quantitative measures and can therefore be less bias indicators of customer experience, when compared to explicit or qualitative methods (Alves et al., 2012). This leads to the next section, explaining why consumers' responses are sometimes biased.

Biases in Consumer Responses

A user's perception towards a system is commonly measured through self-reported measures. However, it is difficult for users to precisely report on their own experience as they may be influenced by multiple biases, often unwillingly. Prior research shows that there is an important discrepancy between what users feel during the experience and how they recall it afterwards (Cockburn et al., 2017; Eich & Schooler, 2000). Research suggests that retrospective evaluations are often biased and that human memory is influenced by peak moments (Kahneman et al., 1993; Kahneman et al., 1997; Redelmeier & Kahneman, 1996). According to Fredrickson and Kahneman (1993) (p. 46), *"[...] most moments of an episode are assigned zero weight in the evaluation and a few select "snapshots" receive larger weights".* This means that those snapshots are usually the only things remembered from a previous moment. Two examples of the snapshots are described by the peak effect and the peak-end rule. The peak effect is that the user tends to better remember the most intense moment of the experience, while the peak-end rule is that the user's impression about the experience tends to be influenced by its final moment (Cockburn et al., 2015). Therefore, when asked about remembering a precise moment, users can lack confidence because of both the process of remembering and the act of evaluation (Kahneman et al., 1997).

While remembering can be difficult because of the loss over time of the ability to recall certain details of the context, the time spent between the experience and the moment of recalling can also impact the biases related to the operation of remembering (Eich & Schooler, 2000; Kahneman et al., 1997). Furthermore, it has been shown that the intensity of the emotions felt plays an important part in the recalling process and that negative memories tend to be better remembered than positive ones (Ariely, 1998; Baumeister et al., 2001). As for the act of evaluation, is has been shown that hedonic and utilitarian moments are remembered differently, but that both are influenced by effects that could bias their retrospective evaluation (Langer et al., 2005). Therefore, current methods used to measure and map user experience may be subject to multiple biases, as they are based on human retrospective evaluation. Using implicit physiological measures is a potential way to get around those biases.

Advantages and Disadvantages of Using Psychophysiological Measures

Over the years, many physiological and psychophysiological measures have been developed to evaluate users' responses such as electrocardiography (ECG), respiration rate, skin-based measures (EDA), blood pressure, ocular measures and brain measures (EEG) (Charles & Nixon, 2019). With the increase popularity of e-commerce, it has become necessary to take into account users' emotions when interacting with an interface, as users' decisions are often based on hedonic motivations rather than utilitarian ones (Bradley & Lafleur, 2016). However, research shows that interrupting users during a task negatively affects their affective states, therefore biasing results (Bailey et al., 2006). Hence, to improve human-computer interaction in e-commerce without interfering with the interaction, using physiological measures can be extremely useful (Dufresne et al., 2010). In domains such as entertainment technologies, physiological measures are far most robust in finding differences between participants and tasks than current subjective methods (Mandryk et al., 2006). Another advantage is that data is collected in real-time, which allows to precisely identify peaks without relying on user's memory. For example, a study on mental workload on air traffic controller operations showed that using real-time eye movement data allowed for deeper insights that subjective ratings might not have discovered, therefore allowing designers to detect problems earlier in the design process (Ahlstrom & Friedman-Berg, 2006). Moreover, capturing data can often interfere with the validity of the results, as users can be obstructed or distracted by the settings or methods used. Using unobtrusive tools to capture psychophysiological data allows users to use a given technology in a realistic way, giving more reliable insights while reducing biases of explicit measures as well, as it can be used in a complimentary way to give more validity to the results (de Guinea et al., 2009).

Furthermore, using implicit psychophysiological measures allows to test multiple new factors that can not be accurately reported by the users at any given moment in time. Many of those measures are constructs related to user experience, such as valence, arousal, and cognitive load (de Guinea et al., 2009). For example, a study on equipment installation found that success was negatively impacted by the level of the user's arousal (Nah & Xiao, 2018, p. 372). This result could not have been found with the same accuracy without the help of psychophysiological implicit measures. Another recent study, where users were asked to retrospectively review their previous interaction with a website at every moment in time, show that the user's accuracy of the evaluation of their previous emotions was extremely low and often completely incorrect (Nah& Xiao, p.132), therefore showing the utility of more accurate measures.

Although physiological measures open new ways for researchers to understand user behavior, it also comes with some disadvantages. Since it is a relatively new area of application, definitions and ways of measuring physiological constructs such as workload or arousal often vary between studies. This makes it difficult to compare results across studies, and to replicate and validate findings (Charles & Nixon, 2019). Also, physiological measurement tools can be sensitive to extraneous noise, further complicating comparison between studies. For example, EDA, a skin-based measure of change in electrodermal activity varies with temperature, level of humidity, time of the day and season, which are all difficult to control (Kramer, 1990). Furthermore, in some cases, participants do not react the same way in a laboratory setting as in a real-life setting. For example, a study measuring mental workload for plane pilots showed that measures taken during a real flight tasks were completely different from measures taken during the same task done in a laboratory setting (Wilson, 1993). Another study using cardiovascular responses also showed a weak correlation between laboratory and real-world contexts (Johnston et al., 1990). Moreover, recent research has shown that a single measure is not sufficient to satisfy validity requirements and therefore, triangulation is necessary in order to obtain valid results (Charles & Nixon, 2019). Triangulation is also necessary because a same physiological reaction can be elicited for different reasons, depending on the context and the user's previous experiences (Charles & Nixon, 2019). Basically, triangulation allows for better data interpretation and therefore more useful insights.

Method

We collected data in order to identify psychophysiological pain points during an e-commerce interaction. The goal was to develop a method that would accurately identify pain points and combine them with eye-tracking data in order to gain key insights into the user journey. A second goal was also to compare the pain points identified using psychophysiological data with the ones identified qualitatively by users retrospectively. To increase the generalization of our results, we used three different websites in order to obtain different sources of pain points. This allowed us to compare pain points both across websites and between participants using the same websites.

Context

We used online grocery shopping as the study context. This context has numerous advantages. First, it involves high complexity arithmetic tasks for multiple items, as users need to figure how much they need of each product (Desrochers et al., 2015). This need for multiple items forces the customer to accomplish multiple tasks in a single visit and choose between a vast product assortment, which makes a session longer than a traditional e-commerce session, even if more convenient than a trip to the grocery store (Morganosky et al., 1999). Second, online grocery shopping also generates risk as users need to trust the website regarding both the freshness and the quality of products as well as confidential data such as credit card and phone information (Citrin et al., 2003). This lack of trust can cause pain points, as users are already potentially opposed to buying fresh products online or filling out their personal information, making them more sensitive towards potential problems. Third, online grocery shopping is an uncommon or unusual transaction for users, as in 2016, only 21% of consumers globally have already bought fresh online groceries (Nielson, 2017). Finally, in this specific context, consumers were more involved in the task as they were buying their own groceries rather than having a simulated goal, compared to other studies where the nature of the task is artificial (Nah & Xiao, 2018, p. 327).

Design, Sample and Procedure

Twenty-one students and young professionals (mean age: 23) were recruited via our institution's panel and were divided between three equal groups of seven participants, each group shopping on a different online grocery website. Using three different websites allowed to illustrate possible comparisons between websites as well as determine if the results were generalizable. Participants had one task: they were asked to do their grocery shopping online, buying items they really needed and paying using their own credit card to maximize ecological validity. The task was the same for all three groups. It lasted between forty-five minutes and an hour, excluding the baseline measures. Participants had to spend at least 50$ and were asked to select the store where they would go to pick up their order in the following days. They had to buy at least one fruit, one vegetable and one piece of meat to make sure they would navigate sufficiently on the website. Participants had to fill out a questionnaire before the task, right after the task, and after picking up their order from the store. An interview was also conducted right after the task by an experienced moderator, in order to know qualitatively how the user felt about the task. In that interview, the user was specifically asked about the positive and negative aspects of his online grocery shopping experience. Every participant received a $60 cash compensation to reimburse their groceries. Each participant completed a consent form beforehand and this project has been approved by the Institutional Review Board (IRB) of our institution.

Measures

During the interaction with the assigned website, non-intrusive tools were used to capture the users' reactions in real time. A Tobii X-60 eye-tracker (Stockholm, Sweden) sampled at 60 Hz, as suggested by Laeng et al. (2012) was used to capture eye-tracking data, and Tobii Studio was used to record the experience. The use of eye tracking data allowed to identify precisely where the participant was looking at every second and the recording allowed to go back afterwards, without interfering with the interaction. Arousal was measured using electrodermal activity (EDA) with the Acqknowledge software (BIOPAC, Goleta, USA). EDA is a precise indication of physiological arousal and its variation throughout time (Hassenzahl et al., 2003). Emotional valence was measured using facial emotion recognition with the FaceReader™ software (Noldus, Wageningen, Netherlands). FaceReader™ was used to observe facial movements to calculate emotional valence, from negative to positive (Den Uyl & Van Kuilenburg, 2005). The Observer XT (Noldus, Wageningen Netherlands) software was also used to synchronize apparatus and event markers.

At the end of the experiment, a qualitative interview was conducted with each participant, where users we asked explicitly about the positive and negative aspects of the task, in order to verify what pain points were noticed by the participants. Qualitative data was analysed using Reframer from Optimal Workshop to find trends between participants. This was done in order to compare the added value of the implicit and explicit measures in the construction of the journey map.

Calculations of pain points using a specific threshold was done using statistical software SAS 9.4 and results were then illustrated as a journey map using Tableau®. In this particular context, to be qualified as a pain point, the data point needed to be both in the ninetieth percentile of EDA (i.e., high arousal) and in the tenth percentile of valence (i.e., large negative valence). Each pain point was validated manually using the time code of the recording in Tobii Studio. It was also used to put markers at the beginning and ending of each subtask, in order to color code them in Tableau®.

These tools allowed to identify and label the emotional peaks. In sum, the visualization method allowed us to accurately and precisely identify the psychophysiological pain points using non-intrusive tools and ensure that our insights were representative of what the users really felt by comparing the results of the quantitative data (implicit pain points) with the qualitative data (explicit pain points).

Results

Our results show that the temporal occurrence of psychophysical pain points can be accurately identified. Using a journey map representation, the evolution of valence (y axis) and arousal (size of dot) over time (x axis), was sampled for every single second (see Figure 2). We color-coded each subtask, i.e., shopping, account creation, payment, time selection, and store selection to better visualize the order as well as the number of times the participant came back to that subtask. As an optimized journey is expected to be linear (i.e., no coming back to a previous subtask), this allows us to see where potential problems could be as well. For example, in Figure 2, we can see that the participant started with shopping, then switched to time and store selection, before returning to the shopping task. He then returned to time selection, before proceeding with account creation and payment. As the journey is relatively linear, there are not many pain points along the way. Pain points were identified using a different shape and colour as the other dots, in other to distinguish them. Pain points are illustrated by red squares (Figure 2) and calculated using a specific threshold. In Figure 2, most pain points are toward the end of the interaction, in the payment and account creation subtask, except for one located in the shopping subtask. It can also be noticed that some pain points are successive, as they come one second after the other. We called those pain periods, as they usually have the same source. For example, there is a pain period labeled « Enters his last name ». This means that this specific task was painful for successive seconds, therefore showing the importance of improving this specific task compared to other single pain points. Finally, the visualization method allows to add labels to the online user journey, to identify the reason behind the pain points visually, so that with one glance, one can understand what is wrong for a specific participant. For the participant below (Fig. 2), we can see that the experience was relatively painless until the end, where s/he experienced many pain points during the payment and account creation subtasks, mainly when entering personal information, such as first name, last name, postal code, and credit card information.

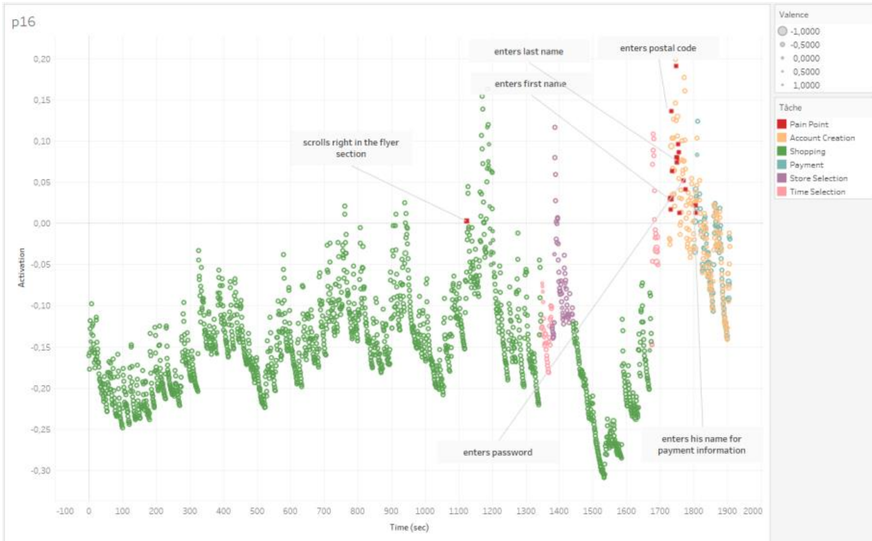


Fig. 2. Visualization of the online user journey for one participant

Furthermore, the visual representation of the user journey allows for an easier comparison between participants. This allows to compare the duration of consumer journeys, as well as the order and duration of the different subtasks and the location of the pain points. In the example below (Fig. 3), one can see that the 6th participant took more than twice the time of the 1st one to complete the same task. All participants started with the shopping subtask, probably because it is the most intuitive way to start. The 2nd participant made his store selection early in the process and that did not cause any pain points, compared to three other participants, that did the same subtask later and experienced pain points doing it. A possible reason explaining those results could be that choosing your store at the beginning shows you only the food items available at the store chosen. If you chose later, some of the products in your cart could become unavailable, causing pain points to the participants because they either had to find a substitute or delete the item from their cart. This method can also be used to compare journeys between different companies. For example, comparing how many pain points were related to shopping or payment for different competitors is a good way to benchmark how well the company is performing in different areas.

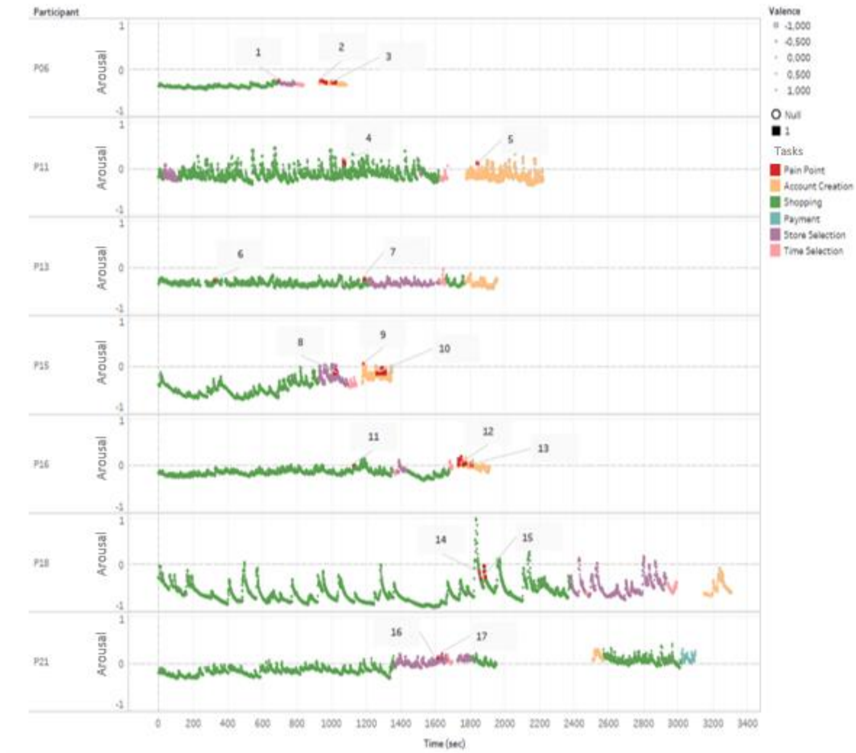


Fig. 3. Comparison of Different Participants. Legend: Numbers indicate the pain points number.

Labelling those pain points also allowed us to compare the experience truly felt by the participants with that users mentioned afterwards. This gave additional insights by identifying pain points that were not mentioned qualitatively by the participants afterwards but most importantly, showed us specific moments when the participant clearly mentioned that a specific subtask went well, while the pain points identified clearly showed otherwise. For example, Fig. 4 shows that the participant reported that he had no problems filling out his credit card information. However, its body reactions showed otherwise.

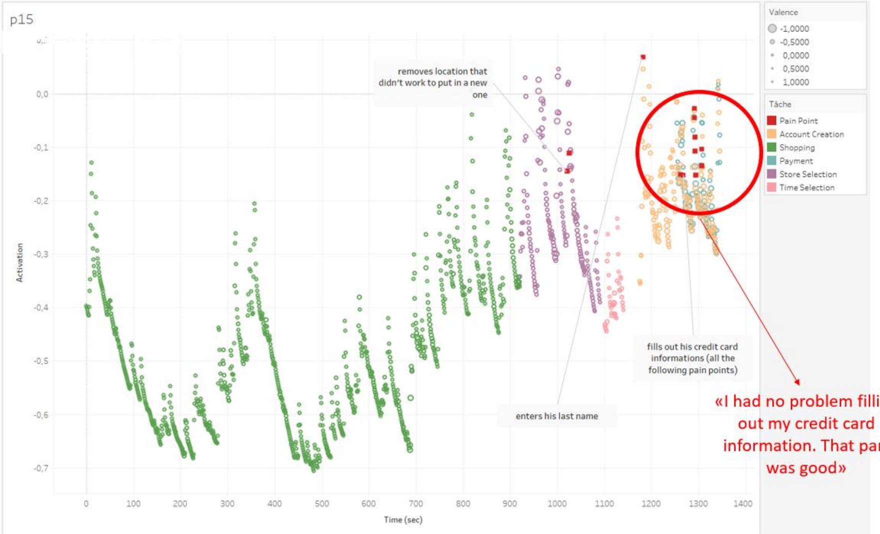


Fig. 4. Comparison of Qualitative and Quantitative Data for One Participant: Pain Points not Identified

Our results showed that less than 25% of pain points were identified qualitatively by the participant afterwards. Out of the 65 pain points or pain periods identified for the 21 participants, only 16 were mentioned as a negative point afterwards (24,6%). Most surprisingly, 5 out of those 65 pain points were clearly mentioned as specifically positive by the participants, while the physiological data clearly showed otherwise, as you can see in Fig. 3 below. Results between grocers are surprisingly similar and are shown in Table **1.** Details of Pain Points per Grocer 1.

Table 1. Details of Pain Points per Grocer

	Grocer 1	Grocer 2	Grocer 3	Global
# of total pain points	72	47	43	162
# groups of pain points	17	27	21	65
Pain points identified verbally	4/17 (24%)	7/27 (25,9%)	5/21 (23,8%)	16 24,6%
Pain points mentioned as strengths by user	2/17 (12%)	0/27 (0%)	3/21 (14%)	5 8,7%

Discussion and Concluding Comments

Our results show that the temporal occurrence of implicit psychophysical pain points can be accurately identified and that the visual representation of the user journey allows for an easier comparison between participants. Moreover, results showed that less that 25% of pain points were identified qualitatively by the participant afterwards and that some pain points were clearly mentioned as specifically positive by the participants, while the physiological data clearly showed otherwise.

This study contributes to the existing user experience literature by proposing a reliable method to visualise peak emotional reactions experienced by users while performing a task. Thus, providing more precision and reliability in identifying pain points when compared to pain points mentioned by users after the task (Fang et al., 2014). It also introduces the notion of implicit psychophysiological pain points, which, compared to explicit pain points previously used in the literature, allows to identify more pain paints and gives more reliable insights by potentially reducing biases of explicit measures (de Guinea et al., 2014; Nah & Xiao, 2018, p. 327).

The results also have managerial implications. First, prior work by Georges et al. (2017) explained the importance of several factors when developing new UX evaluation tools using physiological measures, such as the ability to locate issues, the ease of use, and the reduction of the analysis time. This new method allows both practitioners and researchers to identify psychophysiological pain points easily and the visualization allows to interpret and analyze more efficiently the results. This study contributes to user experience's evaluation tools by using physiological data to assess how users truly felt during an online task, providing more precision and accuracy in identifying pain points when compared to pain points mentioned by users after the task. Therefore, if practitioners are interested in identifying pain points in order to improve interfaces, implicit pain points provide a more comprehensive list. However, if practitioners are interested in what users remember or think of their interface (e.g. attitude), explicit pain points should be used. Second, this study clearly shows that without the implicit emotional measures of users, it would have been extremely difficult to identify pain points, showing the relevance of this current study. Moreover, in an online grocery shopping context, pain points need to be identified in a much more precise way. The new visualization method presented in this study acknowledges this need, so companies can not only identify the «painful» steps, but the exact moment the pain point happened. Moreover, this new method is useful to benchmark user experience across interfaces, which can be used in prototype comparisons or competing interfaces.

Furthermore, some limitations need to be acknowledged. First, this visualization has so far only been applied to an online grocery shopping context and has not been tested in a hedonic context or a context that has a lot of arousal variations. Secondly, the experiment was about forty-five minutes to an hour long, excluding the baseline measures. This can be a limitation, as participants could have gotten tired and the pain points found in the final parts could be related to participants' fatigue rather than actual problems with the interaction. Finally, as they were only 7 participants per grocery website, this was not a large-scale study, mostly due to the high cost of obtaining the data. Hence, additional studies in different contexts as well as of different duration of time and with a greater number of participants could help with the generalization of these results.

In conclusion, using this new visualization method allows to identify implicit psychophysiological pain points in the user's experience, by targeting moments when the user had both a high level of arousal and a negative valence, compared to his baseline state, which meant that s/he felt an intense negative emotion. Identifying those pain points and combining them with eye-tracking data gives key insights into the online user journey and helps identify common negative moments between users. It also allows to gain a deeper understanding of the pain points that participants failed to identify during the post-task interview as well as compare the experience felt by the participants, either between tasks or between companies.

References

Ahlstrom, U., Friedman-Berg, F.: Using eye movement activity as a correlate of cognitive workload. International Journal of Industrial Ergonomics. 36, 623-636 (2006)

Alves, R., Lim, V., Niforatos, E., Chen, M., Karapanos, E., Nunes, N.: Augmenting Customer Journey Maps with quantitative empirical data: a case on EEG and eye tracking, https://arxiv.org/abs/1209.3155.

Ariely, D.: Combining experiences over time: the effects of duration, intensity changes and on-line measurements on retrospective pain evaluations. Journal of Behavioral Decision Making. 11, 19-45 (1998)

Bailey, B. P., Adamczyk, P. D., Chang, T. Y., Chilson, N. A.: A framework for specifying and monitoring user tasks. Computers in Human Behavior, 22(4), 709-732 (2006)

Baumeister, R., Bratslavsky, E., Finkenauer, C., Vohs, K.: Bad is stronger than good. Review of General Psychology. 5, 323-370 (2001)

Bettencourt, L. A., Ulwick, A. W.: The customer centred innovation map. Harvard Bus. Rev. 10(1), pp. 93 (2008)

Bradley, G., LaFleur, E.: Toward the development of hedonic-utilitarian measures of retail service. Journal of Retailing and Consumer Services. 32, 60-66 (2016)

Charles, R., Nixon, J.: Measuring mental workload using physiological measures: A systematic review. Applied Ergonomics. 74, 221-232 (2019)

Citrin, A., Stem, D., Spangenberg, E., Clark, M.: Consumer need for tactile input. Journal of Business Research. 56, 915-922 (2003)

Cockburn A, Quinn P, Gutwin C.: Examining the Peak-End Effects of Subjective Experience. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp. 357–366 (2015)

Cockburn, A., Quinn, P., Gutwin, C.: The effects of interaction sequencing on user experience and preference. International Journal of Human-Computer Studies. 108, 89-104 (2017)

Den Uyl, M.J., Van Kuilenburg, H.: The FaceReader: online facial expression recognition. In: Proceedings of Measuring Behavior, pp. 589–590 (2005)

Desrocher, C., Léger, P. M., Sénécal, S., Pagé, S. A., Mirhoseini, S.: The influence of product type, mathematical complexity, and visual attention on the attitude toward the website: The case of online grocery shopping. In: SIGHCI 2015 Proceedings. 16. (2015)

de Guinea, A., Markus, M.: Why Break the Habit of a Lifetime? Rethinking the Roles of Intention, Habit, and Emotion in Continuing Information Technology Use. MIS Q. 33, 433 (2009)

de Guinea, A., Titah, R., Léger, P.: Explicit and Implicit Antecedents of Users' Behavioral Beliefs in Information Systems: A Neuropsychological Investigation. Journal of Management Information Systems. 30, 179-210 (2014)

ISO: ISO 9241-210:2010. Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems (formerly known as 13407). International Organization for Standardization (ISO), Switzerland (2010)

Dufresne, A., Courtemanche, F., Tep, S. P., Senecal, S.: Physiological measures, eye tracking and task analysis to track user reactions in user generated content. In: 7th International Conference on Methods and Techniques in Behavioral Research, p. 218 (2010)

Eich E., Schooler JW.: Cognition/emotion interactions. In: E. Eich, J. F. Kihlstrom, G. H. Bower, J. P. Forgas & P. M. Niedenthal (Eds) Cognition and emotion Oxford University Press 3 29 (2000)

Fang, Y., Qureshi, I., Sun, H., McCole, P., Ramsey, E., Lim, K.: Trust, satisfaction, and online repurchase intention: The moderating role of perceived effectiveness of e-commerce institutional mechanisms. MIS Quarterly. 38(2), 407-427 (2014)

Georges, V., Courtemanche, F., Sénécal, S., Léger, P.-M., Nacke, L., Pourchon, R.: The adoption of physiological measures as an evaluation tool in UX. In: Nah, F.F.-H. (eds.) HCIBGO 2017. LNCS, vol. 10293, pp. 90–98. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58481-2_8

Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. Behav. Inf. Technol. 25(2), 91–97 (2006)

Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messungwahrgenommener hedonischer und pragmatischer Qualität. In: Szwillus, G., Ziegler, J. (eds.) Mensch & Computer 2003, vol. 57, pp. 187–196. Vieweg+Teubner Verlag, Wiesbaden (2003). https://doi.org/10.1007/978-3-322-80058-9_19

Forrester Research: Mapping the Customer Journey (2010)

F. F.-H. Nah and B. S. Xiao (Eds.): HCIBGO 2018, LNCS 10923, pp. 132-144, 2018. https://doi.org/10.1007/978-3-319-91716-0_1

F. F.-H. Nah and B. S. Xiao (Eds.): HCIBGO 2018, LNCS 10923, pp. 327-337, 2018. https://doi.org/10.1007/978-3-319-91716-0_1

F. F.-H. Nah and B. S. Xiao (Eds.): HCIBGO 2018, LNCS 10923, pp. 372-381, 2018. https://doi.org/10.1007/978-3-319-91716-0_1

Johnston, D., Anastasiades, P., Wood, C.: The Relationship Between Cardiovascular Responses in the Laboratory and in the Field. Psychophysiology. 27, 34-44 (1990)

Kahneman, D., Fredrickson, B., Schreiber, C., Redelmeier, D.: When More Pain Is Preferred to Less: Adding a Better End. Psychological Science. 4, 401-405 (1993)

Kahneman, D., Wakker, P., Sarin, R.: Back to Bentham? Explorations of Experienced Utility. The Quarterly Journal of Economics. 112, 375-406 (1997)

Kramer, A.: Physiological Metrics of Mental Workload: A Review of Recent Progress. Defense Technical Information Center, Ft. Belvoir (1990)

Laeng, B., Sirois, S., Gredebäck, G.: Pupillometry. Perspectives on Psychological Science. 7, 18-27 (2012)

Lallemand, C., Gronier, G.: Méthodes de design UX. Eyrolles, Paris (2018)

Langer, T., Sarin, R., Weber, M.: The retrospective evaluation of payment sequences: duration neglect and peak-and-end effects. Journal of Economic Behavior & Organization. 58, 157-175 (2005)

Mandryk, R., Inkpen, K., Calvert, T.: Using psychophysiological techniques to measure user experience with entertainment technologies. Behaviour & Information Technology. 25, 141-158 (2006)

Mangiaracina, R., Brugnoli, G., Perego, A.: The eCommerce Customer Journey: A Model to Assess and Compare the User Experience of the eCommerce Websites. Journal of Internet Banking and Commerce, 14(3) (2009)

Mirhoseini S.M.M., Léger PM., Sénécal S.: The Influence of Task Characteristics on Multiple Objective and Subjective Cognitive Load Measures. In: Davis F., Riedl R., vom Brocke J., Léger PM., Randolph A. (eds) Information Systems and Neuroscience. Lecture Notes in Information Systems and Organisation, vol 16. Springer, Cham (2017)

Moon, H., Han, S., Chun, J., Hong, S.: A Design Process for a Customer Journey Map: A Case Study on Mobile Services. Human Factors and Ergonomics in Manufacturing & Service Industries. 26, 501-514 (2016)

Morganosky, M., Cude, B.: Consumer response to online grocery shopping. International Journal of Retail & Distribution Management. 28, 17-26 (2000)

Mylopoulos, J., Chung, L., Yu, E.: From object-oriented to goal-oriented requirements analysis. Communications of the ACM. 42, 31-37 (1999)

Nenonen, S., Rasila, H., Junnonen, J. M., Kärnä, S.: Customer Journey–a method to investigate user experience. In: Proceedings of the Euro FM Conference Manchester pp. 54-63 (2008)

Nielson: What's in store for online grocery shopping- Omnichannel strategies to reach crossovers shoppers (2017)

Rawson, A., Duncan, E., Jones, C.: The truth about customer experience. Harvard Bus. Rev., 91(9), pp. 90-98 (2013)

Redelmeier, D., Kahneman, D.: Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. Pain. 66, 3-8 (1996)

Richardson, A.: Using customer journey maps to improve customer experience. Harvard Bus. Rev., 15(1), pp. 2-5 (2010)

Rosenbaum, M., Otalora, M., Ramírez, G.: How to create a realistic customer journey map. Business Horizons. 60, 143-150 (2017)

Teixeira, J., Patrício, L., Nunes, N., Nóbrega, L., Fisk, R., Constantine, L.: Customer experience modeling: from customer experience to service design. Journal of Service Management. 23, 362-376 (2012)

Wang, B., Miao, Y., Zhao, H., Jin, J., Chen, Y.: A biclustering-based method for market segmentation using customer pain points. Engineering Applications of Artificial Intelligence. 47, 101-109 (2016)

Wilson, G.: Air-to-ground training missions: a psychophysiological workload analysis. Ergonomics. 36, 1071-1087 (1993)

Chapitre 3 : Deuxième Article

Identifying and Visualizing Psychophysiological Pain Points: A Step-by-Step Method[2]

Caroline Giroux-Huppé[1], Sylvain Sénécal[1], Marc Fredette[1], Shang Lin Chen[1], Bertrand Demolin[1], Pierre-Majorique Léger[1]

[1] HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, Québec, Canada H3T 2A7

INTRODUCTION

An implicit psychophysiological pain point can be defined as a precise moment in time, when the user both feels a high level of emotional arousal and negative emotional valence, compared with his baseline state. It can be identified using psychophysiological measures, allowing us to build a metric using a specific threshold. An explicit pain point, however, is more commonly used in business or marketing research in order to identify touchpoints generating negative emotions in users. It can be identified using qualitative data, by analyzing words spoken by users during a task (Wang et al., 2016). It can be argued that both the implicit and explicit methods have their strengths and can be used complementary. However, retrospective explicit measures currently used in research lack temporal precision and are influenced by multiple biases, such as peak effects (Cockburn et al., 2017; Ariely, 1998) and emotional ones, such as negative emotions that tend to be remembered more easily than positive ones (Baumeister et al., 2001). Prior research also shows that using explicit measures alone can produce inaccurate responses (Ortiz de Guinea et al., 2013). Recent research suggests that current methods comparing user journeys are far from flawless, at it assumes all touchpoints are equally important for all users, which is far from the truth (Rosembaum et al., 2017). As for implicit measures of

[2] Chapitre de livre en préparation pour soumission

pain points, UX literature shows there is a lack of proper methods to accurately identify implicit psychophysiological pain points.

Accurately identifying the temporal occurrence of psychophysiological pain points that are often not conscious or not recalled by the participant can have a huge impact for UX designers. By targeting those precise moments and identifying the potential reasons behind the pain, companies could fix problems on their websites or mobile applications. Those insights can be crucial in the development of better interfaces that would allow for a better user experience. Comparing a company's pain points with its competitors or another version of their interface can also be extremely useful to benchmark their performance, improve their weaknesses, and leverage their strengths.

The current approaches are mostly qualitative, using focus groups or interviews, or a mix of qualitative and quantitative, using self-reported questionnaires, but all rely only on the users' retrospective assessment of their own emotions (Mucz et al., 2018; Wang et al., 2016; Lee, 2014). Recent methods using pain paints found them by using common keywords or sentences analysis (Wang et al., 2016). A prior study used pain points based on user's dissatisfaction to segment customers into different groups (Lee, 2014). However, prior research in user experience show that there is a growing demand for data-driven recommendations, requiring the use of quantitative research (Georges et al., 2017). Even if a user is asked about a given task right after it is completed, prior research shows that the difference between the actual experience and what is recalled can be quite different (Eich & Schooler, 2000; Cockburn et al., 2017). Furthermore, users can forget multiple aspects, positive or negative, of the task as well as not even being aware of feeling certain emotions. Thus, when being asked about negative aspects of an experience that can be relatively long, it is unrealistic to expect the user to remember accurately all the pain points felt, thus losing valuable insights.

We propose a systemic approach that allows researchers to identify pain points more accurately, from a temporal and content standpoint. By identifying all pain points, i.e., that users failed or not to identify when asked post hoc, this approach gives additional insights on the online user journey and allows a deeper understanding of users' emotions. In the next pages, we will explain the experiment conducted as well as the step-by-step method used to calculate, visualise, identify, and compare psychophysiological pain points. We will then explain the benefits of the method compared to traditional ones as well as the tools and software needed to replicate it.

WHAT WAS THE EXPERIMENT?

To illustrate our proposed method, online grocery shopping was used as the study context. Online grocery was an ideal context for the experiment, because it is a relatively new online shopping context for consumers that could generate risk, as people want to make sure they get the right product, in the right quantity, as the right level of freshness (Huang and Oppewal, 2006). It could also involve complex arithmetic tasks, as consumers need to calculate the right quantities of many unpacked items (Desrochers et al., 2015). Moreover, online grocery shopping is still uncommon, as only 21% of global users had bought fresh groceries online in 2016 (Nielsen, 2017). Twenty-one (21) participants were recruited using our institution's panel, mainly composed of students. It was a between-subject experiment, participants were randomly assigned to one of three different online grocery shopping websites. Using three different websites allowed to illustrate possible comparisons between websites as well as determine if the results were generalizable. As illustrated in Figure 1, the task was the same for all three groups: They had to shop for their groceries online, with the specific instructions of buying products according to their weekly needs, using their own credit card to maximize ecological validity. Participants had to spend at least $50 and were compensated with $60, to reimburse their groceries. They then had to pick up their groceries in the following days at the local store of their choice (that part of the journey is not included in the visualization).

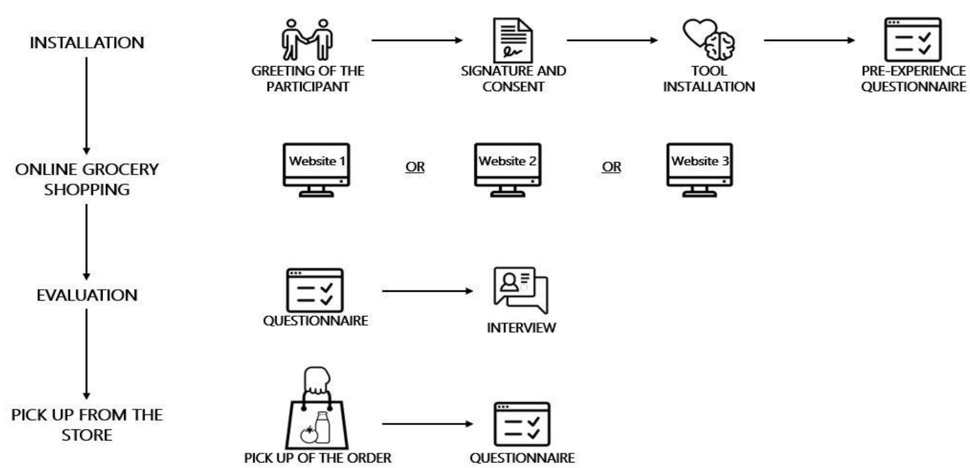


Figure 1: Experiment Design

As no precise instructions were given about the navigation on websites, it allowed to observe the user's real online journey. They could all start with what felt more intuitive to them, whether it was shopping, store selection, or account creation and then continue as they wanted. The three websites all had similar shopping subtasks, except one that additionally had email confirmation as a subtask.

Data was collected in a laboratory setting, with advanced technological tools that were managed by trained research assistants. Eyetracking, using a Tobii Pro X-60 eye tracker (Stockholm, Sweden), electrodermal activity, using the AcqKnowledge software (Biopac, Goleta, USA), facial emotion recognition, using the FaceReader™ software (Noldus, Wageningen, Netherlands), and a session audio-video recording device (Tobii Studio) were used to collected data. Several pretests were performed to ensure the quality of the collected data.

The AcqKnowledge software (Biopac, Goleta, USA) was used to clean EDA data and extract an arousal metric and the FaceReader™ software (Noldus, Wageningen, Netherlands) was used to analyse facial expressions and calculate the emotional valence. In addition, markers were added to identify the beginning and end of the different subtasks using Tobii Studio (e.g. create an account, shop for products). Data from each system (Biopac, Facereader, Tobii) was then converted into text files that were uploaded and processed using CUBE HX (Montreal, Canada). Once processed, the data was available in a large text file containing one line per millisecond with a column per metric for all tools (pupil size for Tobii, valence from FaceReader,, etc.). This file contains all the information needed for pain point identification!

PAIN POINTS: HOW TO FIND THEM?

Step 1: How to calculate pain points?

Pain points are calculated using a specific threshold. In this particular context, the following formula was used:

$$\textbf{(1) Pain Point} = \text{EDA} > \text{EDA}_{p90} \textbf{ AND } \text{Valence} < \text{Valence}_{p10}$$

This means that to be qualified as a pain point, participants need to be both in the ninetieth percentile of EDA (i.e., high arousal) and in the tenth percentile of valence (i.e., large negative valence). To choose the threshold, we built upon a psychophysiological study by Lewinski et al., (2014) that used the top 10% peak values of facial expressions of emotions felt by consumers during advertisements to evaluate their emotions. The result is a column of 0s and 1s, the former if it does not qualify for a pain point and the latter if it does. The formula makes sense, considering that a participant needs to have a high level of arousal while feeling a negative emotion to feel a psychophysiological pain point. The 10% rule used in that case is context-based. The goal is to get a relatively manageable average of pain points per participant (e.g., 10), regardless of the time spent on the task, as having too many pain points would make them less valuable for prioritizing features to improve on the interface. To determine the threshold that would best work for any experience, the best way would be to start with the 90/10 approach as presented above (ninetieth percentile of EDA and in the tenth percentile of valence) and see how many pain points are generated on average. If there are still too many points, we suggest changing the threshold to find a combination that fits the experiment context.

Before manipulating the data, histograms and scatter plots were used to check for obvious outliers that could affect the results' validity. There are no specific criteria for outliers, as it can vary depending on your data set. In this dataset, the rule used was the following:

$$\textbf{(2)}\ \mathbf{EDA_{outlier} > EDA\_mean + 2\ standard\ deviations}$$

Using this rule, no outliers were identified. In this study, data was not normalized, it was rescaled to (-1 to+1) and no baseline was used. However, it does not matter if there is one. Adjusting for baseline does not affect the quantities of pain points found or which pain points are found, as a baseline consists of subtracting a constant, therefore not changing with data points are the highest. As for normalization, the same rule applies. Rescaling as no effect on the relative position of the data points.

Step 2: How to visualise pain points?

The journey map representation can be done using any visualisation or analytical software. As it represents a journey, time needs to be on the x axis, with preferably one data point per second. If the experiment is very long, it can be regrouped to allow a one-page visualisation, but we do not recommend it as you lose precision. Then for the y axis and size of the circles, you might have a choice of variables to choose from such as EDA, EKG, valence, and pupil dilation. If you are conducting exploratory research, you should choose the two most discriminant variables. Then, out of those two, the one with the highest variance should go on the y axis, as it makes it easier to observe variability. The remaining variable can be represented as the size of the circle, for every data point (See Figure 2). In the current case (Figure 2), the time, in seconds, is represented on the x axis, arousal on the y axis and valence using the size of the circles. To clearly identify pain points, they should have a different color and shape than the rest of the dots in order to distinguish them. In this study, they were represented as red squares, compared to round dots of various colors for each subtask (See Figure 2). There were between 5 and 7 subtasks, depending on the website used by participants. Shopping, time selection, store selection, payment, and account creation were common to all websites. Two of them also had order verification and one had email confirmation. It is important to understand that participants did not know that there were different subtasks. They were created by analyzing user journeys and finding common subtasks between the three grocery websites. They were then separated with markers in the recording software (i.e., Tobii Studio). We color-coded each subtask to visualise the order in which participants performed each subtask as well as the number of times participants came back to a specific subtask. As an optimized user journey should include each subtask only once, in a linear way, this shows where potential problems could lie as well. Is it important that subtask colors are different from the pain points so that the latter can clearly stand out. Finally, the last step is to label the different pain points. This can be done directly into the visualisation tool or added manually on the pictures of the visualisation, with a tool such as PowerPoint.

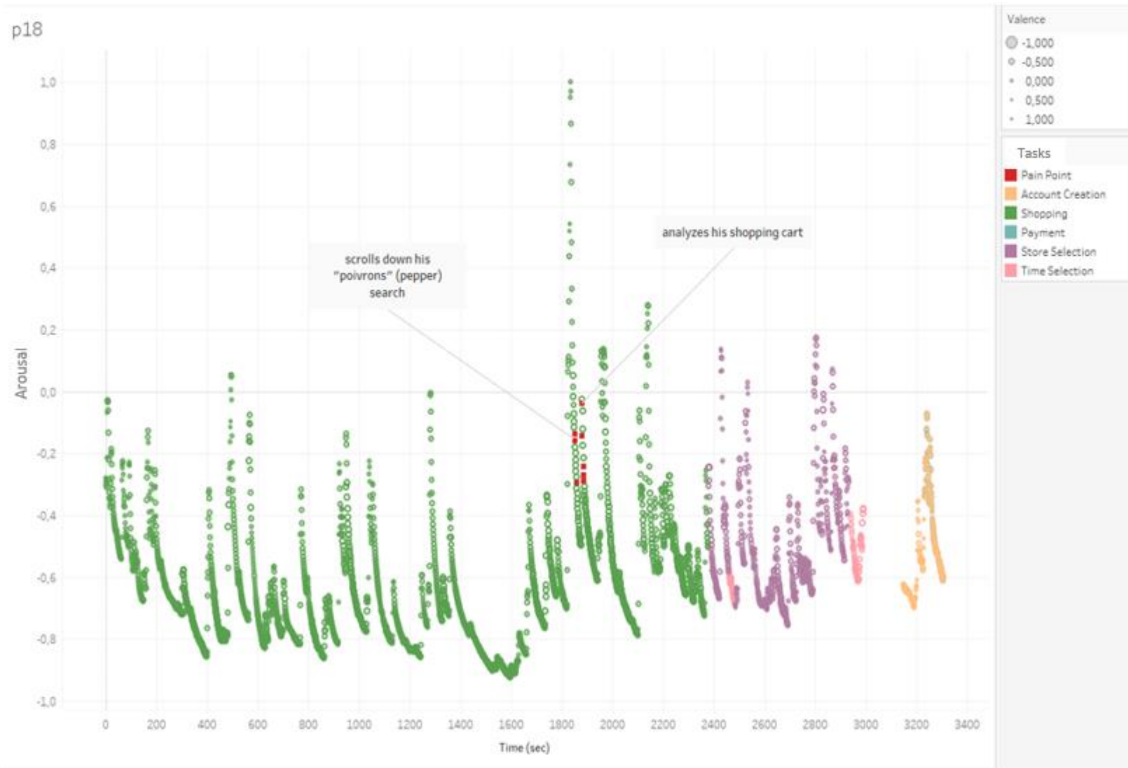


Figure 2: Sample user journey map with pain points labels

You might notice successive pain points (e.g., pain points that are one second after the other). In that case, they are considered different pain points as they are different data points, but are often put together in the interpretation as the reason behind is often the same. They are considered « pain periods » (See Figure 3).

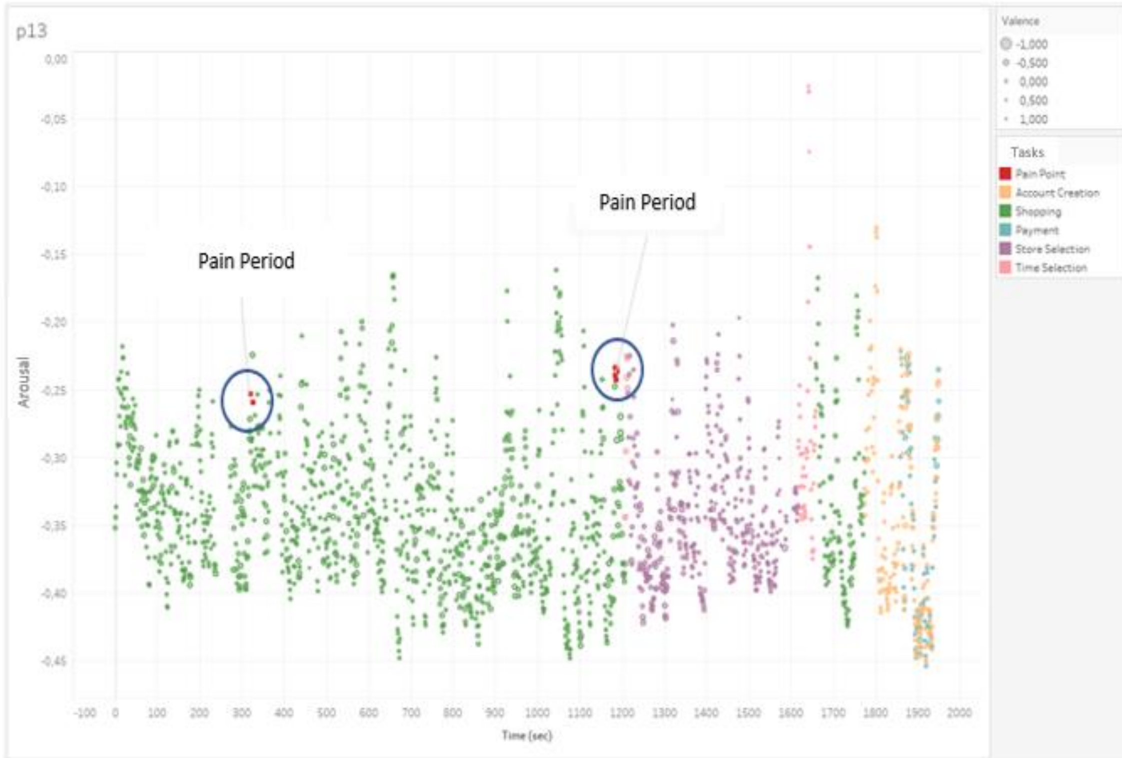


Figure 3: Sample user journey map with pain periods

Step 3: How to manually identify pain points?

In order to label pain points, you need to identify at what time pain points happened in the video recording for each participant. Once the data has been analyzed to identify pain points, you need to filter your previously generated column of 1s and 0s to keep only the «1»s. In the current experiment, this was done using Microsoft Excel, with one tab per participant. This allows to see the number of seconds after the beginning of the first task that the pain point happened. The next step is transforming those numbers into a format easier to find in the video recording software. To do so, we need to transform seconds into (minutes)m:(seconds)s. Here is the formula used:

(3) =CONCATENATE(INT(D2)," m : ",ROUND((D2-INT(D2))*60,0)," s"), where D2 is the cell having the time in seconds.

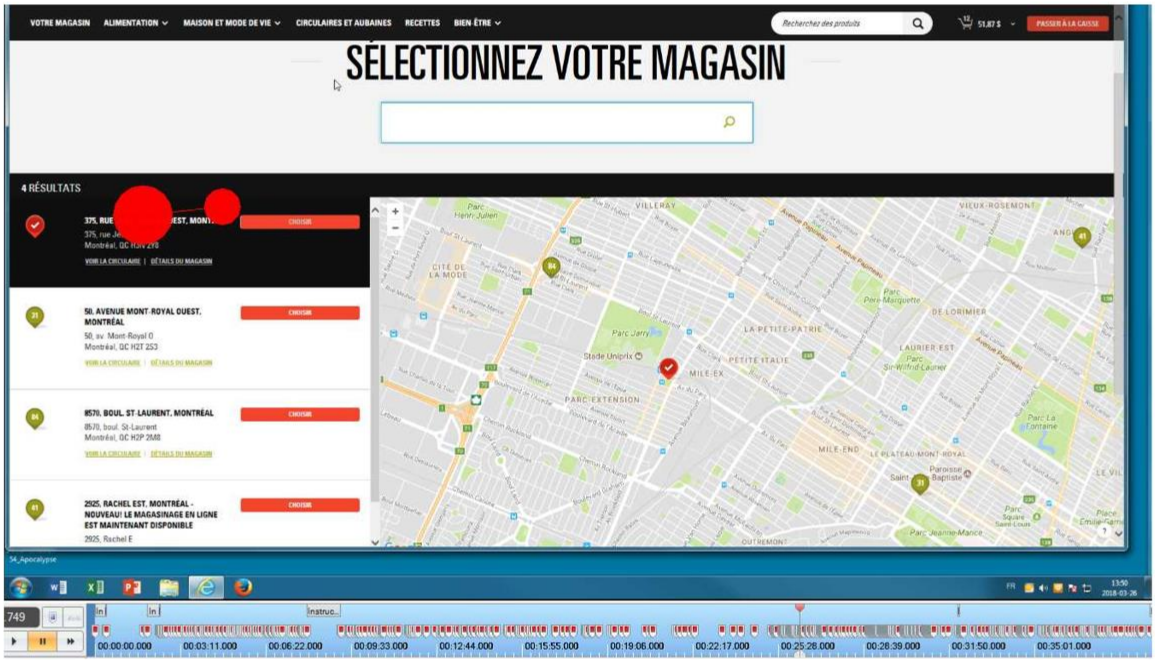
However, any other way to transform seconds into minutes can be applied here.

Then, using the video recording device of your choice, you need to find the exact moment the pain point happened. The use of eyetracking can help you see exactly where the participant was looking at during the pain point, which can be highly valuable and helpful for the identification and interpretation of the pain point. Figuring out the possible reason behind the pain point is not always obvious and requires human judgement. It is important to go back and forth in the minute around the pain point to understand its context. Possible reasons are numerous, but here are some of the most common ones in an online grocery shopping context: use of filters or search bar, not being able to find an item, entering personal information, verifying the order, finding a store. In Figure 4, an example of pain points identified using the formula above is shown, as well as the task the pain points were situated in and the possible reason identified to explain the pain points. To increase pain point identification and interpretation reliability, coding can be done by multiple independent researchers and then assess inter-rater reliability (Gwet, 2014).

Pain Point (1=yes, 0=no)	Time (sec)	Time (minutes:seconds)	Participant	Task Name	Periods	Possible reason behind pain
1	183,91	3m : 4s	P07	Account Creation		Participant enters email address
1	858,81	14m : 19s	P07	Shopping		Search bar: participant looks for what to write
1	916,81	15m : 17s	P07	Shopping		Filters: wine
1	1412,81	23m : 33s	P07	Shopping	pain period	Participant finds dog food in his search for noodles
1	1413,81	23m : 34s	P07	Shopping	pain period	Participant finds dog food in his search for noodles
1	1595,81	26m : 36s	P07	Shopping	pain period	Participant finds dog food in his search for noodles
1	1755,81	29m : 16s	P07	Shopping		Participant clicks on order
1	1770,19	29m : 30s	P07	Store Selection		Participant changes store selection, can't decide

Figure 4: Spreadsheet used to manually identify pain points per participant

In Figure 5 below, you can see the exact moment a pain point has been identified. The red dots represent the gaze of the participant. Moving back and forth can easily be done using the arrows at the bottom left as well as the timeline at the bottom.



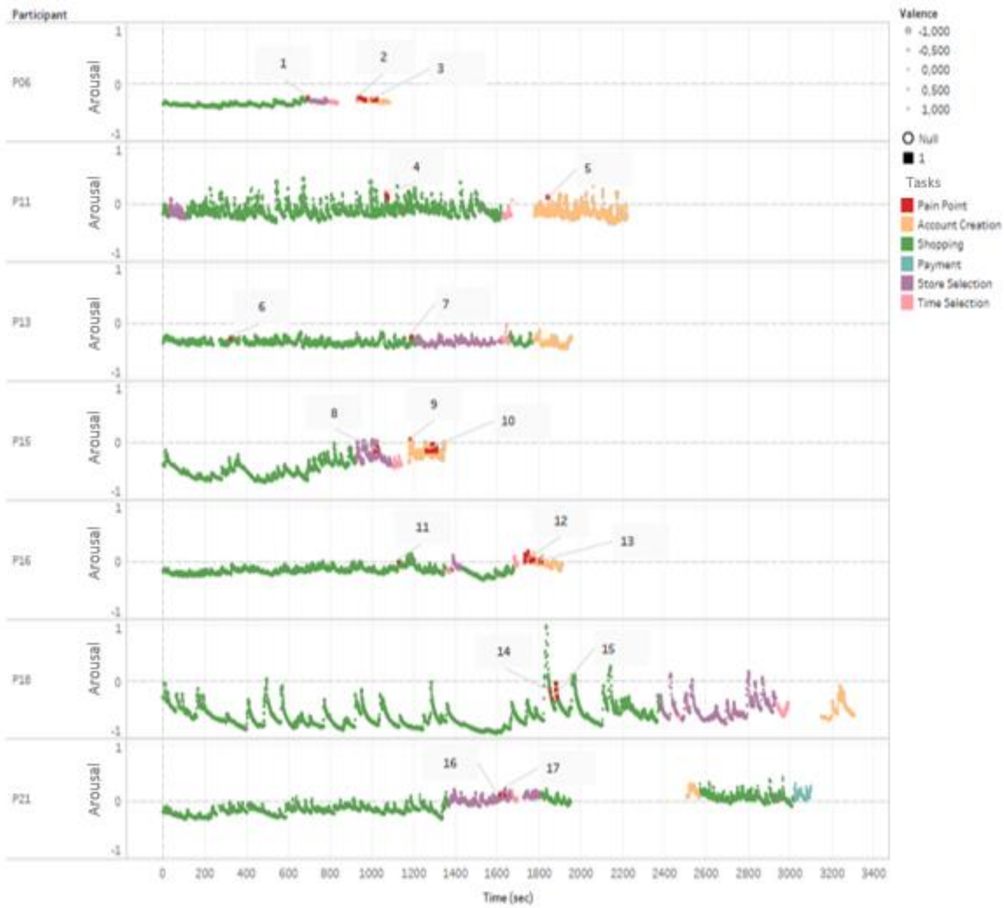
Legend: Red dots represent the gaze of the participant

Figure 5: Example of moment identification using Tobii Studio

Step 4: How to compare participants' journey map and pain points?

We now have seen how to calculate, visualize, and manually identify pain points. However, one of the big advantages of this method is that it allows to compare the different user journeys, both over time and between subtasks.

As you can see in the Figure 6 below, all seven participants shopping at the same grocer are represented on one page. This allows an easy comparison of the time given to each subtask, the number of times the subtasks were performed, the location of pain points in different subtasks and in the overall experiment, as well as the length of every journey.



Legend: Numbers (from 1 to 17) identify pain points

Figure 6: Comparison of users' journey (Numbers represent pain points)

As you can see in Table 1, pain points for all participants were classified by subtask. This allows an easy comparison of the performance of each grocer, for every subtask. In the above example, you can clearly see that Grocer 1 needs to improve its shopping navigation, as 14 of the 25 pain points identified are related to shopping, but also it has a really good payment method, that clearly surpasses its competitors, with no pain points associated to it across the seven users. On the contrary, Grocer 2 clearly needs to improve its payment method as well as its account creation form, as both tasks have 7 pain points related to them out of 18. Grocer 3 has 25% of its pain points related to store selection, which is not excessive, but still higher than its competitors. These examples are only the beginning of what you can observe by comparing experiences. Digging deeper into the shopping experience could help gain additional insights.

Table 1: Comparison of pain points distribution between subtasks, for each grocer

Subtask	**Grocer 1**	**Grocer 2**	**Grocer 3**
Account Creation	3/25 (12%)	7/18 (39%)	5/17 (29%)
Shopping	14/25 (56%)	3/18 (17%)	6/17 (35%)
Payment	0/25 (0%)	7/18 (39%)	2/17 (12%)
Store Selection	2/25 (8%)	1/18 (6%)	4/17 (24%)
Time Selection	1/25 (4%)	0/18 (0%)	0/17 (0%)
Order Verification	1/25 (4%)	0/18 (0%)	0/17 (0%)
Connexion	2/25 (8%)	0/18 (0%)	0/17 (0%)
Email Confirmation	2/25 (8%)	NA	NA

Thus, as you have seen, identifying and interpreting pain points can be done in 4 simple steps:

1) Calculating them using a chosen threshold;
2) Visualizing the user journey using the visualization tool of your choice;
3) Identifying the reasons behind them using the recording of the experiment;
4) Comparing users over time and tasks or condition.

In the next section, we will look at the benefits of using this approach instead of the ones traditionally used.

SO WHAT?

In UX research, comparing users' experience is usually done using retrospective explicit methods, either qualitative (i.e., interviews) or quantitative (i.e., questionnaires). These methods may be biased. For instance, prior research suggests that users tend to remember more the final part of a task (i.e., peak-end rule, Cockburn et al., 2015). However, the main problem with these methods is all the relevant moments that failed to be identified by users. To verify our assumptions regarding users' unawareness or their own emotions or unable to recall them accurately, we compared the psychophysiological pain points identified using our method with the pain points mentioned by the participants in a qualitative interview done right after the shopping task, with a question asking them precisely what the negative aspects of the experience were. Interviews were conducted by an experienced moderator and analyzed using Optimal Workshop's *Reframer* tool.

Comparing Users' Psychophysiological Experiences with Qualitative Data

Using Grocer 3 as an example, we identified a total of 72 pain points. Those 72 pain points were grouped into 17 pain periods. Out of those 17 pain periods, only 4 were explicitly mentioned by participants, which represents less than 25% of the real pain points. Even more surprisingly, 2 participants explicitly mentioned that a certain subtask was positive, when the visualization clearly shows they were experiencing pain points they were not aware of (see Figure 7).

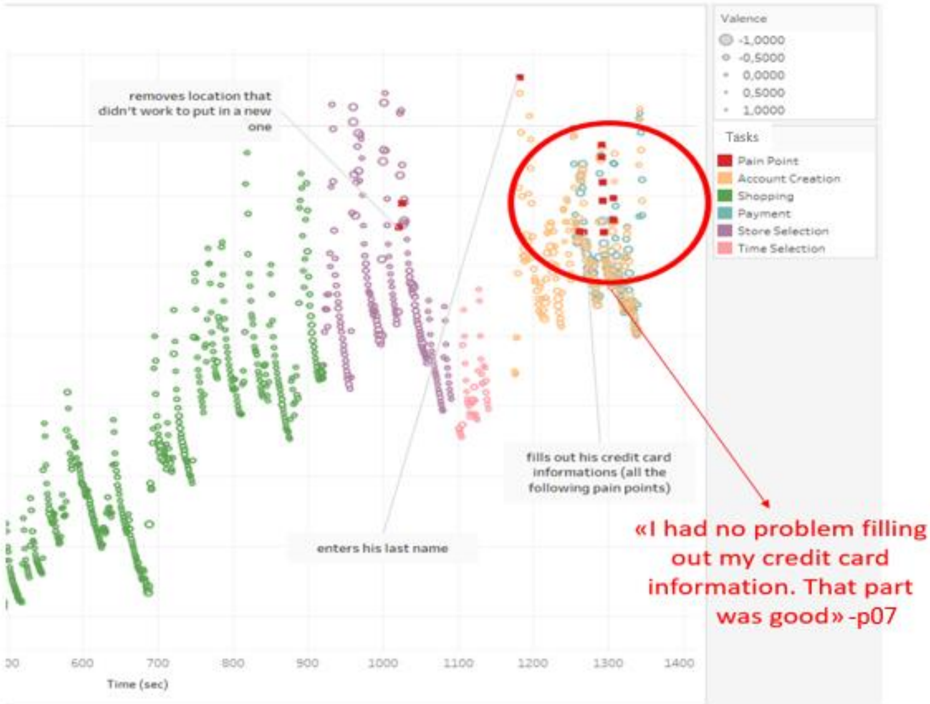


Figure 7: Difference between pain points identified implicitly and mentioned by a participant in the post-task interview.

Furthermore, as shown in Table 2, results were relatively similar across grocers, except for the lack of contradiction by users interacting with the Grocer 2 website. Therefore, this suggests this method could potentially be generalizable to other contexts. A future work could conduct a similar study with the goal of generalizing the use of pain points.

Table 2: Comparison of pain points identification between grocers

	Grocer 1	Grocer 2	Grocer 3	Global
Pain points identified verbally	24%	25,9%	23,8%	24,6%
Pain points mentioned as strengths by user	12%	0%	14%	8,7%

Thus, our results demonstrate that there is a deeper level of understanding that can be reached by researchers using the visualization of psychophysiological pain points. For companies, gaining insights into users' real experience can be highly valuable as it can allow UX designers to adapt their interface to offer a better user experience.

Now that we have shown how using that technique can benefit both academic and corporate domains, we will take a closer look at the tools necessary to replicate the experiment as well as different issues to consider.

WHAT DO YOU NEED TO REPLICATE THIS EXPERIMENT?

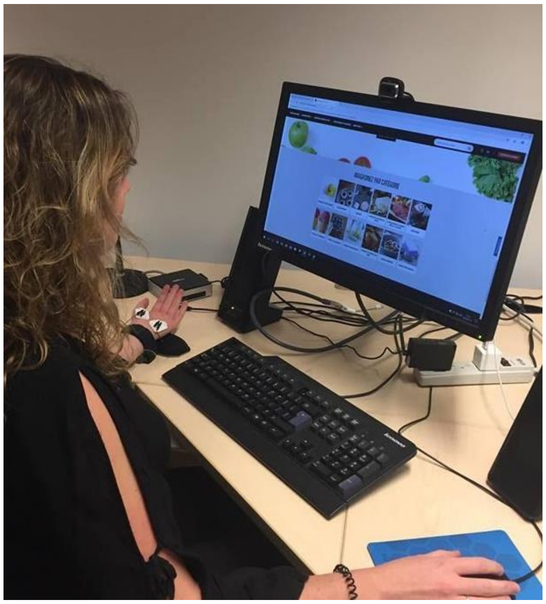
In order to replicate the experiment, a good experimental design is needed, as well as different tools and software to capture the data in a non-intrusive way. The experiment needs to be conducted in a laboratory, where exterior factors are controlled for and researchers are trained on the usage of the different data collection tools.

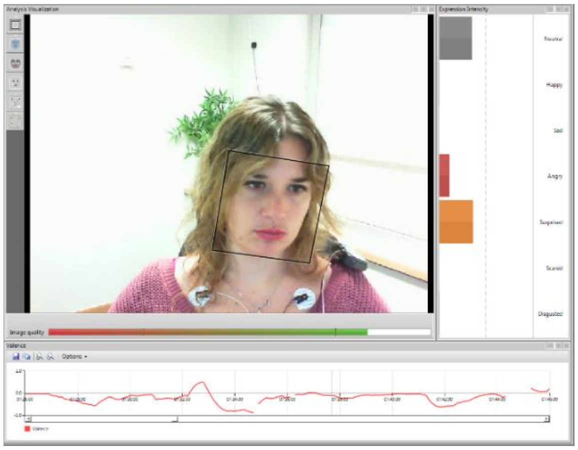
What tools are needed?

An **eye-tracking device** is needed to capture eye-movement data as well as pupil dilation. Precisely, it also allows to manually identify the pain points using the video recording. For this study, a Tobii Pro X-60 eye tracker (Stockholm, Sweden) was used, sampled at 60 Hz, as suggested by Laeng et al. (2012) and Tobii Studio was used to record the experience.

Figure 8: A Tobii Pro X-60 eye tracker (Stockholm, Sweden)

A **facial recognition software** is needed to analyze users' emotions. In order to identify psychophysiological pain points, a measure of the user's emotional valence in real time is necessary. For this study, facial emotion recognition was performed with the FaceReader™ software (Noldus, Wageningen, Netherlands).





Legend: Set up (Left) and Analysis using the software (Right)

Figure 9: FaceReader™(Noldus, Wageningen, Netherlands), a facial emotion recognition software

A precise **measure of arousal** needs to be calculated as well. For this study, arousal was measured using electrodermal activity with the AcqKnowledge software (Biopac, Goleta, USA).

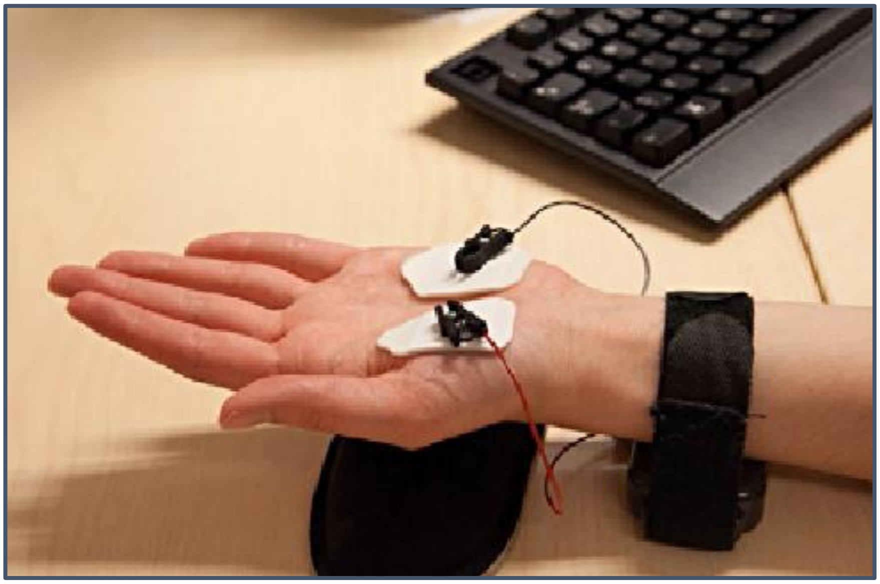


Figure 10: Measure of electrodermal activity using the AcqKnowledge software (Biopac, Goleta, USA)

To identify psychophysiological pain points, user's valence and arousal need to be combined for every second of the experiment. Therefore, a tool is needed to synchronize the data coming from the different software. In this study, Observer XT (Noldus, Wageningen Netherlands) software was used to synchronize recordings and event markers. The data is then triangulated using the Cube UX software, which generates a global data set with all the data.

Once the data is triangulated, it needs to be manipulated. We chose to use the statistical software SAS (Cary, USA), but the calculations could be done using any data manipulation software such as SPSS (IBM, Armonk, USA), Tableau® (Seattle, USA) or even Microsoft Excel (Microsoft, Redmond, USA). There are no advantages to using Excel as this point, as it will take longer, but the manipulation is so simple that any tool at your disposition, even Microsoft Excel, can do the job. There is no need for a statistical software, as no statistical tests were performed. Another option is to use the Tableau software at this point, so everything can be done using Tableau. However, you need to know how to code using Tableau.

Afterwards, results need to be presented using a journey map representation, in order to gain a deeper understanding of the reasons why pain points occur. The representation also allows for a better comparison between users. A visualisation software is needed to represent users' journeys. For the purpose of this study, we used the Tableau® software. Power BI could be used as well, but it seems less flexible than Tableau® and annotations are more difficult to add as you need to add a text box out of the graphic. Furthermore, it is impossible to import SAS dataset into PowerBI, so data has to be converted into an Excel .txt or .csv file beforehand.

The last two steps (statistical analysis and visualisation) could have been done using only Tableau® or Excel. Tableau® is recommended because it is very user-friendly and manipulations can be done easily. Below, in Table 3, is a recapitulation of the tools used at every step.

Table 3: Tools used for every step of data collection, cleaning, and analysis

Steps	Tools used
1. Data collection	
1.1 Eye tracking	Tobii Pro X-60 eye tracker (Stockholm, Sweden)
1.2 Facial emotion recognition (valence)	FaceReader™ software (Noldus, Wageningen, Netherlands).
1.3 Arousal	AcqKnowledge software (Biopac, Goleta, USA).
2. Data synchronization	Observer XT (Noldus, Wageningen Netherlands)
3. Data triangulation	Cube HX
4. Data analysis	SAS (Cary, USA) or SPSS or Tableau® or any data processing software
5. Data visualisation	Tableau®

Thus, in order to be able to replicate this method you need the following tools and software: an eye-tracker, a facial recognition software, a measure of the level of arousal software, a synchronization software, a triangulation software, a data manipulation software as well as a visualization software.

What else should you consider?

As the cost of the experiment quickly adds up (e.g., equipment, salaries, participant recruitment and compensation), it is crucial to make sure you get it right on the first try.

Therefore, it is highly recommended to pretest the whole experiment to make sure that all the data is exploitable, and the synchronization is correctly done.

Also, it is important to consider that although it is a systemic approach, there is still a human factor in order to identify and interpret psychophysiological pain points. Thus, it is important to keep that in mind and make sure to be constant throughout the analysis. One way to do so is to make sure the same person is responsible for the analysis of all the participants or that intercoder reliability is satisfactory and that this analysis is done over a relatively short period of time.

CONCLUSION

In conclusion, we have proposed a systemic approach allowing researchers and UX designers to accurately identify implicit pain points, from a temporal and content standpoint. The proposed method, explained step-by-step, shows how to calculate visualise, identify, and compare implicit psychophysiological pain points. Accurately identifying the temporal occurrence of psychophysiological pain points that are often not conscious or not recalled by the participant can have a huge impact for both academic and practice purposes. By giving us additional insights on the online user journey, we can obtain a deeper understanding of users' emotions. It can lead to designing better interfaces and being able to compare an interface's pain points with both its competitors or another version of the same interface, allowing researchers and designers to improve interfaces weaknesses and benchmark their performance.

References

Ariely, D. (1998). “Combining experiences over time: The effects of duration, intensity changes, and on-line measurements on retrospective pain evaluations”, Journal of Behavioral Decision Making, 11, 19–45.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C. & Vosh, K.D. (2001) “Bad Is Stronger Than Good”, Review of General Psychology, 5 (4). 323-370.

Cockburn A, Quinn P, Gutwin C (2017) “The effects of interaction sequencing on user experience and preference”, International Journal of Human-Computer Studies, 108, 89–104.

Cockburn A, Quinn P, Gutwin C (2015) “Examining the Peak-End Effects of Subjective Experience”,Chi, 1, 357–366.

Desrocher, C., Léger, P. M., Sénécal, S., Pagé, S. A., & Mirhoseini, S. (2015). “The influence of product type, mathematical complexity, and visual attention on the attitude toward the website: The case of online grocery shopping”. In Fourteenth Pre-ICIS SIG-HCI Workshop, Fort Worth, TX.

Eich, E., & Schooler, J. W. (2000) “Cognition/emotion interactions” In E. Eich, J. F. Kihlstrom, G. H. Bower, J. P. Forgas, & P. M. Niedenthal (Eds.), Cognition and emotion (pp. 3–29). New York: Oxford University Press.

Følstad, A., & Kvale, K. (2018). “Customer journeys: A systematic literature review”, Journal of Service Theory and Practice, 28(2), 196-227. doi:10.1108/JSTP-11-2014-0261

Georges V., Courtemanche F., Sénécal S., Léger PM., Nacke L., Pourchon R. (2017) “The Adoption of Physiological Measures as an Evaluation Tool in UX”. In: Nah FH., Tan CH. (eds) HCI in Business, Government and Organizations. Interacting with Information Systems. HCIBGO 2017. Lecture Notes in Computer Science, vol 10293. Springer, Cham

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC.

Huang, Y., and H. Oppewal. (2006). “Why Consumers Hesitate to Shop Online: An Experimental Choice Analysis of Grocery Shopping and the Role of Delivery Fees.” International Journal of Retail & Distribution Management. 34 (4–5), 334–353.

Laeng, B., Sirois, S., and Gredeback, G., (2012). “Pupillometry: a window to the preconscious?” Perspectives on psychological science 7, 1, 18-27.

Lee, S. (2014). “PIG (Pain Is Good) Strategy: Make Customer Centricity Obsolete and Start a Resource Revolution”, iMatchPoint Limited. Hong Kong.

Lewinski, P.; den Uyl, T. M.; Butler, C. (2014). “Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader”, Journal of Neuroscience, Psychology, and Economics, 7(4), 227.

Moon, H., Han, S. H., Chun, J., & Hong, S. W. (2016). “A design process for a customer journey map: a case study on mobile services”. Human Factors and Ergonomics in Manufacturing & Service Industries, 26(4), 501-514.

Mucz, D., & Gareau-Brennan, C. (2018). “MAPPING THE JOURNEY: Engaging with customers, identifying touch points, and developing recommendations at a public library”. Library Journal, 143(17), 32.

Nielsen (2017). “What’s in store for online grocery shopping- Omnichannel strategies to reach crossovers shoppers”, *Nielson Global Connected Commerce Report*, 7-9.

Ortiz de Guinea, Ana, Ryad Titah and Pierre-Majorique Léger (2013). “Measure for Measure: A two study multi-trait multi-method investigation of construct validity in IS research”, Computers in Human Behavior, 29 (3), 833-844.

Rosenbaum, M. S., Otalora, M. L., & Ramírez, G. C. (2017). “How to create a realistic customer journey map”. Business Horizons, 60(1), 143-150.

Wang, B., Miao, Y., Zhao, H., Jin, J., & Chen, Y. (2016). “A biclustering-based method for market segmentation using customer pain points”. Engineering Applications of Artificial Intelligence, 47, 101-109.

Chapitre 4 : Troisième Article

The Mediating Role of Satisfaction in the Relationship between Customer Experience and Intentions: A Psychophysiological Study of Multichannel Grocery Shopping[3]

Caroline Giroux-Huppé, Sylvain Sénécal, Marc Fredette, and Pierre-Majorique Léger

ABSTRACT

Online grocery shopping is increasing globally, although not as much as other product categories. Although brick-and-mortar stores still seem to be consumer's favorite way to shop for groceries, multichannel is emerging, with online being the fastest growing one. As multichannel retailing is growing, customers get both online and physical experiences in the same transaction process and expect the journey to be seamless, as all channels are important in their satisfaction toward the process and the brand. However, grocers have yet to find the right way to optimize the customer journey, both online and offline. As competition is increasing, it has become crucial to understand how to attract and retain customers in order for grocers to remain profitable. To our knowledge, while many studies have looked at satisfaction to predict consumer behavior, there is a lack of research studying the whole process, from pre-purchase expectations to post fulfilment satisfaction in this context. Since it has been shown that final satisfaction is a main determinant of whether consumers will repurchase or not, it is crucial to understand the process that leads to that final satisfaction. To gain a better understanding, a multichannel experiment was

[3] Article en préparation pour une soumission à l'*International Journal of Electronic Commerce (IJEC)*

conducted with 45 participants, both in a laboratory setting and during the store pick-up, allowing us to include implicit and explicit measures. This study makes several contributions to the e-commerce and marketing fields, such as the mediating role played by consumers' satisfaction in the relationship between their online emotional reactions and their intentions as well as the importance of negative emotions.

KEYWORDS AND PHRASES: Online grocery shopping, satisfaction, e-commerce, website design, security, fulfillment, customer service, emotional responses

Introduction

Online grocery shopping is increasing as the global online grocery market is expected to grow at a compound annual growth rate of 17% between 2017 and 2022 (Forrester, 2018). However, even though consumers have been shown to buy more online in general, online grocery shopping is growing at a much slower pace (Anesbury, Nenycz-Thiel, Dawes, & Kennedy, 2016). In 2016, only 21% of consumers globally had already bought fresh groceries online while in other industries, such as fashion, it was 58% (Nielsen, 2017).

In prior research, many reasons have been proposed to explain that difference. The most obvious one being the need to touch fresh groceries before buying them, to check for size and perishability (Citrin, Stem Jr, Spangenberg, & Clark, 2003). Another reason behind consumers reticence is risk aversion, with both product quality and the guarantee to receive products on time being perceived as uncertain (Huang & Oppewal, 2006). A third reason is that shopping online is a much more complex task than shopping for apparel, for example, because it requires multiple decisions and calculations as consumers need to figure out the right quantities needed. Prior research investigating online grocery shopping suggests that when the ordering process involves more arithmetic calculation (for example when fractions are involved), consumers tend to have a negative attitude toward the website (Desrochers et al., forthcoming).

Although brick-and-mortar stores still seem to be consumers' favorite way to shop for groceries, multichannel grocery shopping is emerging, with online being the fastest growing one (Anesbury et al., 2016). In 2017, 30% of U.S. consumers had tried online grocery shopping at least once (Statista, 2017). On the other hand, 20% of U.S. consumers had tried using store pick-up services for their grocery shopping and 8% of U.S. consumers used it as their main way of purchasing groceries (Statista, 2017). E-commerce is now changing the way retailers do business and having a good e-service proposition is crucial to retaining customers (Liao, Palvia, & Lin, 2006). Nowadays, 43% of U.S. millennials have now used online-only retailers for grocery shopping (The Hartman Group, 2017). Therefore, retailers now must adjust to make sure they meet consumers' expectations, both online and offline, to maintain their brand image and stay competitive against online only retailers such as Amazon (Kahn, Inman, & Verhoef, 2018), who already possessed 26% of online food and beverage market shares in the United States in 2016 (Bloomberg, 2016). Transforming non-online grocery shoppers into satisfied loyal online customers is an opportunity for grocery retailers to remain competitive, as it has been shown that high customer retention is linked to higher profits (H. S. Lee, 2013) and returning customers can be five times more profitable than new ones (Gupta & Kim, 2007). Furthermore, as there is more than $5 trillion at stake globally in the grocery category, it is understandable that online retailers feel the need to improve their performance (Forrester, 2018). Therefore, in order to transform potential customers into established frequent online shoppers, it has become crucial to better understand what influences their satisfaction (Cao, Gruca, & Klemz, 2003).

To our knowledge, while many studies have looked at satisfaction to predict consumers behavior (e.g, Bhattacherjee, 2001), there is a lack of research studying the whole process, from pre-purchase expectations to post fulfilment satisfaction (Dan J. Kim, Ferrin, & Rao, 2009). Since it has been shown that final satisfaction is a main determinant of whether consumers will repurchase or not (Richard L. Oliver, 1993), it is crucial to understand the process that leads to that final satisfaction. Past research by Cao et Gruca & Klemz (2003) also highlights the importance of measuring the ordering and fulfilment stages separately, in order to reduce response-effect biases as well as reflect the delay customer experience between the two stages. Furthermore, recent research shows that implicit responses, often unconscious, can interact with explicit ones in the formation of beliefs (de Guinea, Titah, & Léger, 2014). Therefore, we performed a multichannel study from pre-purchase to post fulfilment with 45 consumers who had never done online grocery shopping before, using explicit self-reported measurements as well as implicit psychophysiological measures.

The objective of this paper is to study the online grocery shopping process from beginning to end, while integrating physiological data during the interaction to measure consumers' automatic and often unconscious reactions susceptible to influence their satisfaction. Furthermore, this paper aims to look at both the mediating role of satisfaction in the relationship between emotions and intentions as well as the direct effect of negative emotions on consumers’ satisfaction and intentions.

This paper makes the following contributions to theory. First, by investigating the whole multichannel process into a single consumer journey, results provide a comprehensive understanding of the formation of customer satisfaction, from pre-purchase expectations to post purchase fulfilment. Second, results suggest that a psychophysiological negative emotional responses influence satisfaction toward the online experience, but also influence consumers’ intentions to repurchase and recommend. Finally, results show that satisfaction mediates the relationship between the psychophysiological emotions and intentions. These findings also provide insights for managers by allowing them to better understand the importance of multichannel satisfaction as well as emotions felt during the interaction in order to achieve overall satisfaction as well as positive intentions.

This paper is organized as follows. It starts with a literature review of the main concepts and the theoretical background, followed by the presentation of the proposed conceptual model. Then, it presents the development and justification of the hypotheses, as well as explains the research methodology and the operationalization of the research variables. Finally, it explains the results and discuss their contributions and implications as well as limitations and research avenues.

Literature Review

The Expectation-Confirmation Theory

The Expectation-Confirmation Theory (ECT) is considered dominant in the consumer satisfaction literature (Dan J Kim, 2012). It states that « customer satisfaction develops from a customer’s comparison of post-purchase evaluation of a product or service with pre-purchase expectations » (Dan J Kim, 2012, p. 220). Also called Expectation-Disconfirmation, Disconfirmation Theory, or Expectancy Disconfirmation (Oliver, 1980), it is widely used in both information systems (IS) and marketing research (Khalifa & Liu, 2002; Richard L Oliver & Winer, 1987). The process first described by Oliver

(1980) is illustrated in Figure 1 and goes as follows. Consumers develop an initial level of expectations based on what they already know about the product or service they want to purchase. Then, consumers buy the product and evaluate how well it performed compared to their initial expectations. This gives them a certain extent of (dis)confirmation, depending if the product or service performs as expected, lower, or higher than initial expectations. This determines whether they are satisfied or not with what they purchased. Finally, that level of satisfaction influences repurchase intentions, as satisfied customers will most likely want to buy products again than dissatisfied ones (Oliver, 1980). It is of interest because it can be used to explain repeat purchases, which is more often than not the case in grocery shopping (Dabholkar, Shepherd, & Thorpe, 2000). The ECT allows to look at the formation of users' satisfaction from a multichannel perspective, dividing the process into different stages (Dan J Kim, 2012). Furthermore, this model has the advantage of looking at both the antecedents and the behavioral consequences of customer satisfaction (E. W. Anderson & Sullivan, 1993). The model's predictive ability is established in multiple contexts such as marketing, e-commerce, and information systems, which are all relevant to online grocery shopping (Dan J Kim, 2012). This model applies to both offline contexts, such as brick-and-mortar stores, and online contexts, such as online grocery website. The only differences are the expectations and the stage of the process at which consumers confirm if their expectations are met.

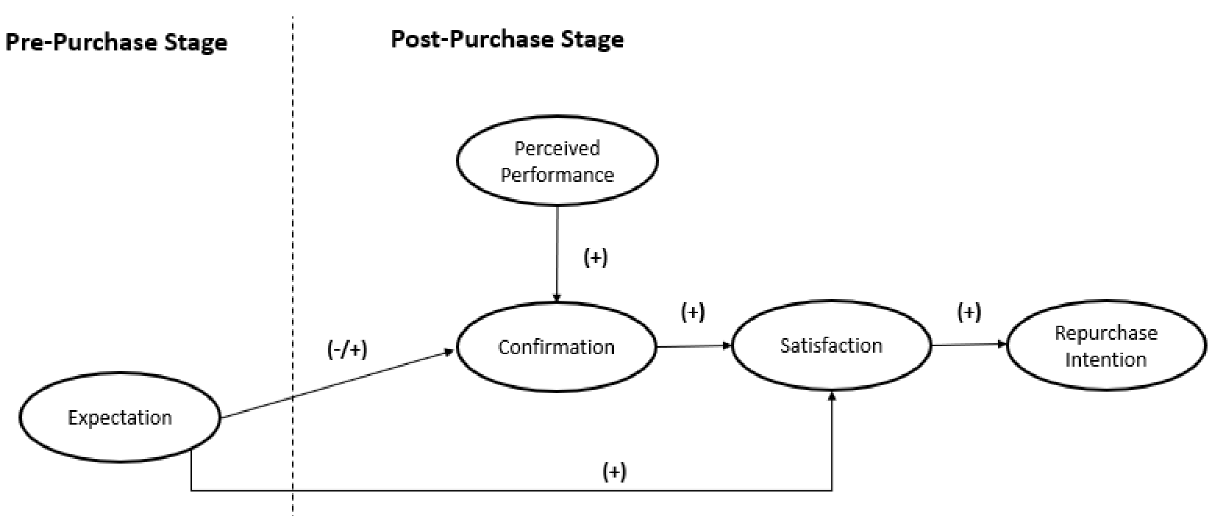


Figure 1. The Expectation-Confirmation Model

However, this model has limits. First, the conceptualization of certain constructs, such as satisfaction and expectations, may vary between studies meaning that studies using the same model do not always refer to the same concepts (Bhattacherjee, 2001). Then, prior critics toward the ECT model claim expectations often differ before and after the experience and therefore, are not a realistic way to predict consumers' subsequent actions (Bhattacherjee, 2001). Most importantly, this model predicts consumer satisfaction using expectations and confirmation. However, those constructs cannot explain by themselves all the variance in consumer satisfaction (Myers, 1991). Most importantly, consumers rely on memory to evaluate confirmation, which has been proven to be subjected to multiple biases such as the peak end rule, that gives a higher importance to the final moment of the experience (Cockburn, Quinn, & Gutwin, 2015, 2017; Schooler & Eich, 2000). In the case of a long transaction such as online grocery shopping, memory can be influenced by multiple biases. Therefore, it is crucial to go beyond perceptions and analyze consumers « real » experience in order to gain a better understanding of consumer satisfaction, such as using the richness of information about user affective state provided by psychophysiological measures (Dirican & Göktürk, 2011).

Measuring Satisfaction Toward Online Service Quality

As e-commerce is increasing, retailers are facing new challenges online and have yet to understand customers' expectations toward the online transactional channel (Yang & Fang, 2004). Customers often evaluate e-commerce service quality as inferior compared to the service received in store (Silvestre, Sue, & Allen, 2009). Hence, evaluating service quality is essential, as it is linked to satisfaction and is a huge part of online retailers' success (Wolfinbarger & Gilly, 2003). Online service quality can be defined as « *overall customer evaluations and judgments regarding the excellence and quality of e-service delivery in the virtual marketplace* » (Santos, 2003).

As the online transactional channel is relatively new and complex, there is no consensus in the literature on how to measure online service quality. Research by Lee & Lin (2005) shows that dimensions of website design, reliability, responsiveness and trust all have a positive influence on overall service quality and customer satisfaction. Similarly, research by Cai & Jun (2003), suggests that four dimensions are determinant of online retailers'

success: website design, trustworthiness, reliable service and communication. Work by Yang & Fang (2004) shows that customers still expect traditional service quality dimensions, such as reliability, responsibility and empathy, even when buying from pure Internet-based suppliers. That same study suggests that integrating Information System (IS) quality dimensions, such as ease of use and security, with traditional service quality dimensions is essential to customer satisfaction (Yang & Peterson, 2004). Finally, a study highlights the fact that most scales focus exclusively on goal-oriented, utilitarian behaviors, forgetting the hedonic quality dimensions (Bauer, Falk & Hammerschmidt, 2006). It suggests adding the online enjoyment dimension, as it influences significantly customers repurchase intentions (Bauer, Falk & Hammerschmidt, 2006).
In previous studies, many scales are used to evaluate online service quality, but the constructs used to evaluate it often have little in common (Loiacono, Watson, & Goodhue, 2002; Yoo & Donthu, 2001; Zeithaml, Parasuraman, & Malhotra, 2002). One of the most commonly used, the WebQual (Loiacono et al., 2002), is more focused on helping web designers design better websites than evaluating the actual service quality felt by consumers, making it less pertinent for service quality measurement (Zeithaml et al., 2002). An older scale, the GAP model Parasuraman et al. (1985) builts on the fact that service quality is a difference between expectation and performance. These gaps are multiple and include the difference between consumers' expectations and management's perceptions of those expectations as well as the difference between consumer's expectation and perceived service (Seth, Deshmukh, & Vrat, 2005). However, while this model is still relevant, it lacks dimensions specific to online service quality. Furthermore, many scales focus exclusively on the interaction with the website interface, while from the consumers' perspective, online shopping includes much more than that with task varying from browsing, ordering, paying, interacting with customer service, picking up the purchase or getting it delivered and finally, evaluating satisfaction toward the purchase (Ha & Stoel, 2009).

The EtailQ is a reliable and valid scale for the measurement of etail quality developed by Wolfinbarger and Gilly (2003). EtailQ is defined as a « conceptualisation and measurement of online retail service quality » (Caruana & Ewing, 2006). A previous study in the IS field shows that most of the etailQ dimensions are strong predictors of customers' satisfaction and allows us to look at both the online transaction and the offline fulfillment part (Li, Aham-Anyanwu, Tevrizci, & Luo, 2015). Moreover, as omnichannel and multichannel retailing is growing, customers get both online and physical experiences in the same transaction process and expect the journey to be seamless, as all channels are important in their satisfaction toward the process and the brand (Piotrowicz & Cuthbertson, 2014). Therefore, this paper integrates the etailQ as a multidimensional construct in the Expectation-Confirmation Theory, in order to gain a better understanding of the distinct role of each dimension in each phase of the process.

Research Hypotheses and Proposed Research Model

This article proposes an adaptation of the Expectation-Confirmation Model that integrates implicit psychophysiological measures as well as the four dimensions of eTail quality proposed by Wolfinbarger and Gilly (2003): website design, security/privacy, fulfillment/reliability and customer service. The proposed research model uses the ECT four times, once for every dimension of eTail quality, in order to be able to measure satisfaction at different moments in time. For example, using the expectation-confirmation theory allows to measure satisfaction right after the purchase, by comparing consumers' expectations right before the purchase to their confirmation right after the purchase. The same goes for the offline part: by measuring consumers' expectations before picking up their groceries from the store and their confirmation after, satisfaction can be measured. In Figure 2, the two online dimensions, website design and security, are regrouped under online satisfaction while the two offline dimensions, fulfillment and customer service, are presented under offline satisfaction.

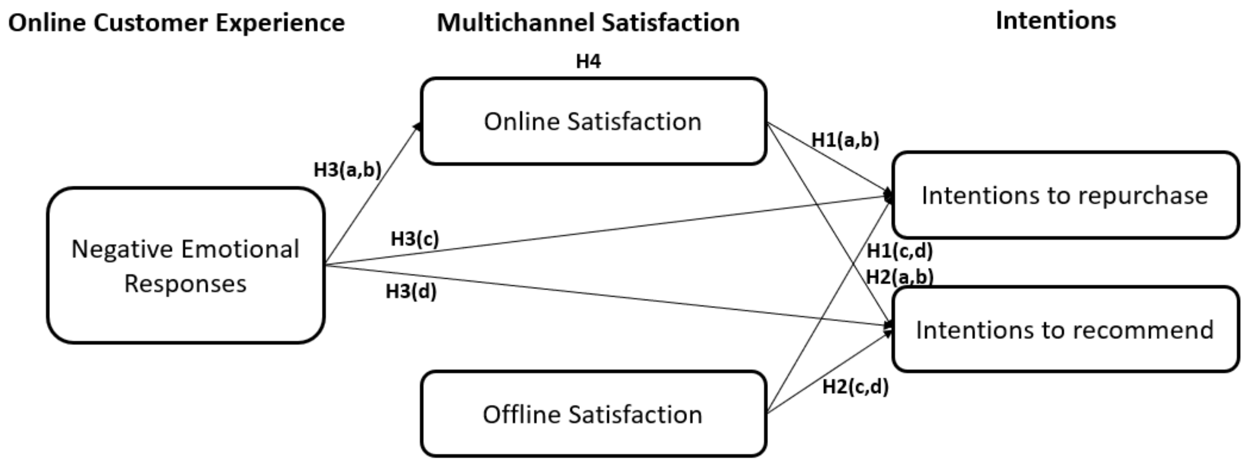


Figure 2. Proposed Research Model

Expectations

In current literature, there is no consensus on how to conceptually define expectations (D. J. Kim et al., 2009). For the purpose of this study, we define expectations as « pre-purchase beliefs or evaluative beliefs about the product » (Richard L Oliver & Winer, 1987, p. 470). Some definitions are broader and define expectations as a state of mind on any subjective subject (Richard L Oliver & Winer, 1987), but we chose that definition for two reasons: It is done before purchase rather than in retrospective, as suggested by Oliver (1981) and it has an evaluative component that is measurable. Expectations can be positive or negative. While most studies chose to focus on the positive side, it is possible to expect a negative outcome and then confirm it afterwards, confirming your beliefs on a negative aspect (Chiu, Hsu, Sun, Lin, & Sun, 2005). While, originally, expectations referred to the performance of a system or product, it is now often used in a much broader sense, as research suggested performance was not always the main concern of consumers (Venkatesh, Thong, Chan, Hu, & Brown, 2011). Accordingly, the consumer's expectations are now looked at through different aspects such as pricing, promotions, product quality, and service quality (E. W. Anderson & Salisbury, 2003). The Unified Theory of Acceptance and Use of a Technology (UTAUT) integrates expectations as performance expectancy to predict behavior. However, it lacks a confirmation/disconfirmation construct that would measure if those expectations were actually met (Venkatesh et al., 2011). Furthermore, expectations evolve over time (Richard L Oliver & Winer, 1987). Prior research on online technologies suggests that expectations are subject to change before and after the first contact with the product, service, or company and that initial expectations have a strong impact on satisfaction (Medina, Rufín, & Rey, 2015). While pre-purchase expectations are often based on word of mouth and mass media, the expectations that follow are directly influenced by their first experience, making them more realistic (Fazio & Zanna, 1981). Expectations also vary between consumers, as people have different baseline levels based on their psychological characteristics as well as the situational context (Bhattacherjee, 2001). Research shows that customer expectations are also influenced by previous experiences as well as their frequency of use, as expectations are expected to adjust and strengthen as the purchase frequency increases (Lin & Lekhawipat, 2016).

Hence, initial expectations are used for comparative judgement (Chiu et al., 2005), making it much harder to satisfy someone with high baseline level expectations (Bhattacherjee, 2001). Moreover, as technology evolves, consumers gain more control over the process and therefore, develop higher expectations toward all phases of the purchase, from the ease of use of the website to the customer service or sales support (Rust & Lemon, 2001)

Confirmation

Confirmation is directly linked to expectations, as it is a direct confirmation or disconfirmation of their initial expectations (Bhattacherjee, 2001). Confirming initial expectations is often easy for consumers, because they get exactly what they expected. However, it can be more difficult when their expectations are not met, as consumers' experience disconfirmation and therefore, a cognitive dissonance. This is coherent with Festinger's cognitive dissonance theory (Festinger, 1957), that suggests disconfirmed perceptions can lead to psychological tension as well as a tendency to distort their perceptions to be more consistent with reality (Bhattacherjee, 2001).

Satisfaction

Satisfaction is a key element when attempting to build strong and successful relationships with consumers, both offline and online. Considering that studies have already demonstrated the importance of satisfaction in increasing loyalty (R. E. Anderson & Srinivasan, 2003), it is crucial to better understand how to influence satisfaction in order to retain customers. In this study, we use Kim, Ferrin & Rao's (2009, p. 237) definition of consumer satisfaction: « *An attitude formed through a mental comparison of the service and product quality that a customer expects to receive from an exchange with the level of quality the consumer perceived after actually having received the service/product* ». Satisfaction is therefore a comparison between the consumers' initial expectations of the overall process and the fulfillment of those expectations (Casaló, Flavián, & Guinalíu, 2008; Cristobal, Flavián, & Guinaliu, 2007). According to Boulding et al., (1993) overall satisfaction is not only multidimensional, but also the aggregation of all previous evaluations of that satisfaction. Since online grocery shopping is a service offered that can expand past the online shopping to the entire consumption experience including the interaction with the brick-and-mortar store's employees, it is crucial to meet consumers' expectations during the whole process. Research has shown that for a consumer to be satisfied, his expectations need to be met or exceeded (Kumar & Oliver, 1997). Furthermore, based on the Expectation-Confirmation Theory (ECT), satisfaction would be a strong predictor of customers' intentions to revisit and repurchase (Lien, Wen, & Wu, 2011; Richard L. Oliver, 1993). Finally, Delone and MacLean, (2003) argue that satisfaction is crucial to consider, as it is a factor that greatly influences technology usage.

Intention to Repurchase & Intention to Recommend

Continuance intentions is the subject of many studies in both the IS and marketing fields. While IS research use models such as the expectation-confirmation theory (ECT) and the technology-acceptance model (TAM) to predict future behaviors, marketing research also focus on recommendation behaviors (Chea & Luo, 2008). In the online shopping context, intention to repurchase can be considered the same as the intention to adopt, as adopting a website means you intend to shop on it again (O. Pappas, G. Pateli, N. Giannakos, &

Chrissikopoulos, 2014). Intention to repurchase is crucial for businesses, since it is well known that long-term relationships between firms and consumers depend heavily on post-purchase decisions (Dan J Kim, 2012; D. J. Kim et al., 2009). Customers make a large proportion of their purchases from a store they already bought from (Richard L Oliver, 1981). Furthermore, multiple studies find that satisfaction influences intentions to repurchase, as well as impacts indirectly through an adjustment of their expectations (Ha & Stoel, 2009; Lin & Lekhawipat, 2014). Some studies suggest that consumer satisfaction is often a stronger predictor of behavioral intentions than more obvious factors, like service quality (Cronin Jr & Taylor, 1994). Additionally, prior research suggests that when looking at the direct relationship between satisfaction and repurchase intentions, the elasticity of repurchase intentions is lower for firms giving their consumers a high level of satisfaction (E. W. Anderson & Sullivan, 1993). Research by Reichheld and Schefter (2000) also shows that satisfaction with the offline process, including fulfillment and customer service, is highly related to customer loyalty, which is similar to intentions to repurchase. As the relationship between satisfaction and intention to repurchase is established in the existing literature (Lien et al., 2011; Richard L. Oliver, 1993), and that satisfaction is shown to be a significant predictor of both continuance and recommendations behaviors (Chea & Luo, 2008), we test both online and offline phases. This leads to the following hypotheses:

H1: Intentions to repurchase are positively influenced by online and offline satisfaction. Specifically, intentions to repurchase are influenced by online website design quality (H1a), online security (H1b), offline fulfilment (H1c), and offline customer service (H1d).

H2: Intentions to recommend are positively influenced by online and offline satisfaction. Specifically, intentions to recommend are influenced by online website design quality (H2a), online security (H2b), offline fulfilment (H2c), and offline customer service (H2d).

Emotions in Customer Experience during Online Grocery Shopping

User experience is a user's perceptions and effects resulting from the use of an interactive system, including his emotions, beliefs, and physiological responses (ISO, 2010). Emotions are an important component of user experience, because consumers' emotional evaluation of that experience allows them to compare possibilities (Russell, 2003) and shape future behavior (Hassenzahl, 2013) . Logically, experiencing a service results in the consumer feeling emotions, that are highly linked with their physiological state (Purves et al., 2001). Therefore, it is important to measure emotions during user experience in order to better understand users' perceptions and behaviors.

Measuring Emotions: Self-Reported vs Implicit Methods

User experience has previously been measured through measurement scales, assessing their past experience. Those explicit measurement scales refer to different aspects of user experience, such as usability (Brooke, 1996; Loiacono et al., 2002), attractiveness (Hassenzahl, Burmester, & Koller, 2003) or pleasure, arousal, and dominance (Bradley & Lang, 1994). However, those measurement scales, even if they ask how consumers feel, usually obtain intellectual and evaluative responses that do not consider the emotional impact, but rather focus on performance, or cognitive aspects (Hartson & Pyla, 2012). As technology is now commonly used in personal aspects of consumers' lives, such as doing online grocery shopping at home, it is becoming crucial to consider emotional responses as well (Hartson & Pyla, 2012).

Furthermore, self-reported methods can be biased, as it is well established that a consumer's instantaneous experience is different from what he actually remembers (Cockburn et al., 2017; Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Schooler & Eich, 2000). Providing a retrospective answer to a question, even a simple one, can often be hard. Different reasons are shown to explain that difference between what really happened and what we felt happened afterwards. First, it is now mainly accepted that bad memories are easier to recall than good ones, just like bad events, bad impressions, and bad interactions (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Therefore, when recalling a long and complex task such as shopping online for groceries, consumers may tend to remember their friction points easier. Then, intensity of the emotions felt and the moment the consumer felt it can also play a role. Ariely (1998)'s work demonstrates that when remembering painful experiences, consumers' judgement of the whole experience was based on intensity during the final moment as well as the latter half of the experiment. This goes accordingly with Cockburn et al.'s (2015) recent findings, that suggests that people's retrospective evaluations are biased by both the final moment (peak-end rule) and the most intense one (peak effect). The tendency to neglect the durations of task during retrospective evaluation also seemed to be a systemic bias (Langer, Sarin, & Weber, 2005). Furthermore, recent research suggests that primary and recency effects have an impact on the accuracy of consumers' self-evaluation of valence and arousal (Lourties et al., 2018). However, work by Walker et al. (1997) demonstrates that, as time pass, consumers' evaluation of the pleasantness or unpleasantness of a moment became less extreme, diminishing even more for unpleasant memories than pleasant ones. Prior research suggests that people often have a hard time discerning their own emotions, as they see them as overlapping experiences rather than isolated states (Posner, Russell, & Peterson, 2005).

Using implicit methods in e-commerce research, such as psychophysiological measures, is a way to reduce those biases, as they are more efficient in finding differences between participants, compared to self-reported subjective methods (de Guinea et al., 2014). Moreover, they allow for real-time data collection, which makes it easier to identify peak moments (Ahlstrom & Friedman-Berg, 2006) (Courtemanche et al., 2017). A recent study by Lourties et al., (2018) shows the usefulness of accurate psychophysiological measures, as the consumers' retrospective subjective evaluations of their previous experience was often incorrect.

Measuring Implicit Negative Emotions: A Combination of Arousal and Valence

As explicit methods to measure emotions might be biased and consumers may have a hard time identifying their emotions, it is of high interest to measure how emotions are felt during the interaction. It is well known that user experience is dynamic and can vary greatly during the task, depending on various factors, such as the complexity of the different subtasks (Wimmer, Wöckl, Leitner, & Tscheligi, 2010). It can also be measured or quantified through many different implicit metrics, such as valence, arousal, or cognitive load (Albert & Tullis, 2013). However, while those metrics may be interesting individually, it is combining them that gives us insights on consumers' emotions while allowing us to pinpoint frustration and confusion peaks (Albert & Tullis, 2013). For example, the circumplex model of affect allows to combine valence and arousal and interpret the results as a particular emotion representing the affective experience (Posner et al., 2005).

Accordingly, building upon previous work by Posner et al., (2005), we combine two constructs that gives valuable insights into consumers' emotions: valence and arousal.

Emotional arousal

Arousal can be defined as a user's emotional state on a spectrum going from quietness to arousal depending on his level of physiological activity (De Guinea, Titah, & Léger, 2013). Few studies have looked at relationships between emotional variables such as arousal and consumer satisfaction (Bigné, Andreu, & Gnoth, 2005). In prior research, arousal is measured with 6 items self-reported scales: cheerful-depressed, quiet-anxious, enthusiastic-calm, nervous-relaxed, active-passive and surprised-indifferent, but it showed a low reliability (Bigné et al., 2005; Kumar & Oliver, 1997). Furthermore, Wirtz et al. (2000) demonstrates that the ideal level of arousal depends on the settings (situation, time, and place) as well as the characteristics of the user. The same study suggests that consumers are sensitive to the «targeted arousal level» they want to reach and that not reaching that target decreases their satisfaction (Wirtz, Mattila, & Tan, 2000). In a utilitarian online grocery shopping process, a high level of arousal might be associated with frustration, but only when combined with negative emotions.

Valence

Valence can be defined as a pleasure-displeasure continuum, with neutral in the middle (Posner et al., 2005). In prior research, valence is measured through self-report measures. A study using a four-item scale that evaluated whether participants felt positively or negatively about a given experience finds that valence directly influences satisfaction (Brady, Voorhees, Cronin Jr, & Bourdeau, 2006). However, psychometrics scales have the disadvantage of only measuring how the user feels at a certain point in time, usually after the task is completed. Some research even implied that user measurement scales sometimes measure the wrong thing as the user may answer differently depending on the success or failure of the task (Zaman & Shrimpton-Smith, 2006). Therefore, for the purpose of this study, valence is measured throughout the experience to assess users' emotions during the task. Work by Ekman (1978, 1984, 1997) suggests that emotions are characterized by patterned change in both expression and physiology and that those changes are distinct for each emotion. Emotions can be assessed using the Facial Action Coding System (Ekman & Friesen, 1978), a « *comprehensive, anatomically based system for measuring all visually discernible facial movemen*t » (Ekman, 1997, p. 12). This measurement system is one of the most intensely used and allows for identifying and coding the intensity of facial movements and emotions (Ekman, 1997). However, Ekman (1992) suggests that reports using a single observer are subjected to multiple bias and errors, highlighting the need for fully automatic face detection and tracking.

Emotions on Satisfaction & Intentions

Traditionally, only cognitive measures were used to explain satisfaction in services, such as the (dis)confirmation of perceived performance (Liljander & Strandvik, 1997). Research now shows that emotions play an important role as well (Liljander & Strandvik, 1997). The role of emotions on satisfaction and behaviors have been studied extensively in the literature, usually in more traditional service settings (Kumar & Oliver, 1997; Mooradian & Olver, 1997). However, limited research has been done regarding the role of emotions in online settings and if so, only consider satisfaction as an emotional variable (Chea & Luo, 2008). It is now well established in the consumer satisfaction literature that attributing negative emotions to a service providers results in a lower satisfaction (Richard L. Oliver, 1993; Taylor & Todd, 1995). For example, studies find evidence of a valence-congruent relationship between satisfaction and negative emotions, where having intense negative emotions lead to a lower satisfaction level (Dube-Rioux, 1990; Hui & Tse, 1996; Richard L. Oliver, 1993; Price, Arnould, & Tierney, 1995). In extended service transactions, where the customer interacts at different points during the service provision, a direct negative relationship is found between negative emotions, such as anger, and satisfaction (Taylor & Todd, 1995). Moreover, multiple studies also find strong relationships between negative emotions and intentions (Inman, Dyer, & Jia, 1997; Zeelenberg & Pieters, 2004). Work by Zeelenberg & Pieters (2004) shows a strong direct relationship between emotions and behavior, that was much stronger than the effect of dissatisfaction on behavior. Research done in a utilitarian service setting suggests that emotions influence both satisfaction and intention to recommend. Surprisingly, it also finds that while positive emotions influence satisfaction even more strongly than negative ones, negative emotions influence more strongly the intention to recommend than positive ones (Rychalski & Hudson, 2017). As it is a utilitarian context, the role of emotions is limited, but they still play a major role in understanding consumers satisfaction (Ladhari, Souiden, & Dufour, 2017). Furthermore, in marketing research, measuring negative emotions can be tricky, as participants might be biased by self-awareness (Meyers-Levy & Malaviya, 1999) and social desirability (Arnold & Feldman, 1981). In advertisement, previous work shows that automatic measures of emotions have higher predictive power than self-reported ones, as they can capture low-order emotions resulting from low-complexity automatic processes (Lewinski, Fransen, & Tan, 2014; Poels & Dewitte, 2006). Hence, this leads us to the following hypothesis:

H3: Negative emotions in the shopping task influence online satisfaction as well as intentions. Specifically, having a more negative experience has a detrimental effect on consumers' satisfaction towards website design (H3a) and security (H3b), as well as intentions to repurchase (H3c), and intentions to recommend (H3d).

Mediation Effects

The mediating role of satisfaction has been shown multiple times (Gustafsson, Johnson, & Roos, 2005). A study exploring antecedents of satisfaction of multichannel customers finds that the indirect effect of perceived multichannel service quality on customer retention is significantly mediated by satisfaction (Hsieh et al., 2012). Another study by Yang and Peterson (2004) shows the mediating role of customer satisfaction in the value-loyalty relationship. Similarly, research by Zboja and Voorhees (2006) finds a mediating effect of satisfaction with the retailer between satisfaction with the brand and repurchase intentions. Moreover, in omni-channel retailing it shows that consumer satisfaction fully mediates the effect between timeliness and loyalty when consumers have to buy online and then pick up in store (Murfield, Boone, Rutner, & Thomas, 2017). Research including emotions suggests that emotional experiences can influence behaviors through satisfaction (Phillips & Baumgartner, 2002). A study by (Han, Back, & Barrett, 2009), shows that satisfaction mediates the effect of emotion factors on intentions to revisit. As, in the current study, we separate satisfaction between online and offline, we propose the following hypothesis:

H4: Online satisfaction mediates the effect between consumers' negative emotions and intentions. Specifically, satisfaction toward the website design mediates the relationship between negative emotions and intentions to repurchase (H4a) and to recommend (H4b) and satisfaction toward website security mediates the relationship between negative emotions and intentions to repurchase (H4c) and intentions to recommend (H4d).

Method

Sample

Forty-five participants, who had never purchased groceries online, were recruited using our institution's participant panel. (M_{age}= 24.1, SD=5,3). Although our distribution is mainly composed of young adults, studies have shown that they represent a very large segment of the total online population, as 43% of U.S. users are aged between 18 and 34 years old (Statista, 2018). In the U.S., users aged between 25 to 34 have the highest monthly average of interest usage amongst all group ages (Statista, 2018). Furthermore, in 2018, 98% of young adults in the U.S. used the internet at least occasionally, compared to 66% for 65+ years old (Pew Research Center, 2018). Participants were composed of 27 women (60%) and 18 men (40%). Although the sample size is small, it is not uncommon for information system studies using psychophysiological measures to have between forty and fifty subjects (Riedl, Fischer, & Léger, 2017). This study was approved by our institution research ethics committee. Each participant received $60 as a compensation, as well as a $10 gift card once they completed the final online questionnaire.

Procedure

Participants were randomly assigned to shop on one of three different online grocery websites and to pick up their order at a grocery store of their choice. We chose to use three websites for external validity purposes. In a pre-test, the level of complexity of the different websites was tested using a panel of ten experts. They had to evaluate the complexity by performing the same shopping task on all websites in a counterbalanced order and rate them using the System Usability Scale (Brooke, 1996). No significant differences were found between the three websites.

Participants had to spend a minimum of $50 and were asked to buy the same items they would usually buy on a shopping trip to the grocery store. They had 45 minutes to complete the task and were verbally informed when they had 15 minutes left. Only three restrictions applied: They had to buy at least one fruit, one vegetable, and one piece of meat. This was asked in order to force participants to explore the different sections of the website. To maximize the ecological validity, participants had to buy their groceries using their own credit card.

As illustrated in Figure 3, the first questionnaire was administered before the participant had the chance to go on the website. Participants were given the name of the website they would later shop on and were asked to answer questions about their expectations toward the online part of the process. The second questionnaire was administered right after participants were done with the online task. Participants were asked to answer questions regarding their satisfaction toward the online part and their expectations toward the end of the process, the offline part. Demographic questions (age, sex, revenue, education, and occupation), as well as questions regarding the participant's lifestyle (online shopping habits, frequency of grocery purchases, monthly amount spent on groceries, distance from the closest grocery store) were also asked. A qualitative interview was also conducted with each participant to highlight the strengths and weaknesses of the online part of the process. The third questionnaire was completed at home by the participants in the 24h after they had collected their groceries at the store, which usually happened between 1 and 5 days after the purchase stage. They were asked to answer questions regarding their satisfaction toward the last 2 dimensions of the EtailQ: fulfillment and customer service (see Table 1). The difference between the participant's answers at this stage and the pre-collect from the store allowed us to compare expectations between actual perceptions, for each item. Participants were also asked questions regarding their intention to repurchase and intention to recommend the grocer.

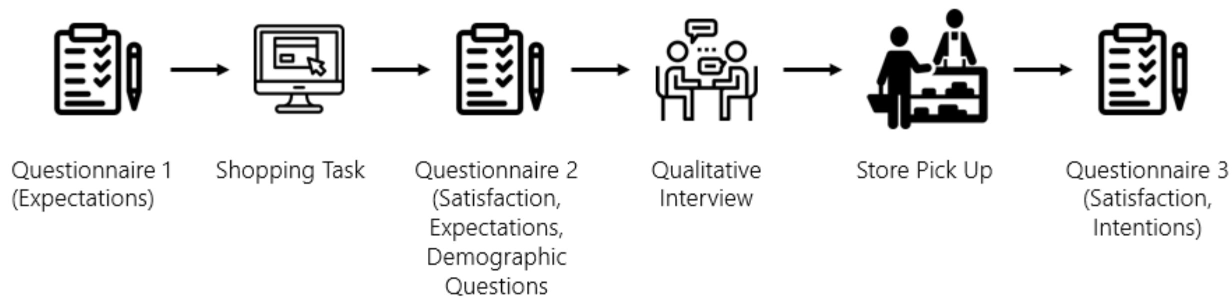


Figure 3: Study Timeline

Instrumentation and Measures

The scale used, the ETailQ, includes four dimensions: website design, privacy/security, fulfillment/reliability and customer service. The website design includes most of the elements of the user's interaction with the interface except customer service and private information. The privacy/security dimension refers to the security related to credit card payments and the privacy of shared information, for example, in the account creation. The fulfillment/reliability dimension can be defined as both the fact that the customer receives what he ordered and that he gets it when expected. Finally, the customer service dimension can be defined as responding to consumers' inquiries quickly in a responsive, helpful and willing way (Wolfinbarger & Gilly, 2003).

The proposed conceptual model required to collect data at four different stages: pre-purchase, during the purchase, post-purchase, and post-fulfilment (after the pick up from the grocery store); allowing us to observe the evolution of the participants' satisfaction as well as the use of different channels. Therefore, three rounds of online measurement scales were administered in a longitudinal design inspired by Kim, Ferrin & Rao (2009)'s research model. This allowed us to measure dimensions that could not have been measured right after, such as customer service and fulfilment, while verifying if expectations change before and after purchase. This addresses prior critics in the literature toward the ECT model, claiming expectations often differ before and after the experience and therefore, are not a realistic way to predict consumers' subsequent actions (Bhattacherjee, 2001). This study used a 14 items scale to measure etail quality, with 7 points Likert-type scales ranging from -3 (Strongly disagree) to +3 (Strongly agree).

Physiological measures were collected during the purchase stage, which was done in laboratory. In the proposed model, we chose to define satisfaction as a comparison between the consumers' initial expectations, of the overall process, and the fulfillment of those expectations (Casaló et al., 2008; Cristobal et al., 2007). Therefore, satisfaction is based on the difference between consumers' expectations and the confirmation of their expectations, asked later in the process. This approach has already been used in the development of a satisfaction model (Khalifa & Liu, 2002).

Explicit Measurement

Expectations were measured using the two online dimensions of a modified version of the EtailQ (8 items), website design and security, where every item was modified to start with «I expect to». One of the very few studies done from beginning to end of the process showed that several dimensions of the etailQ have a significant effect on e-satisfaction (J. Kim, Jin, & Swinney, 2009). As suggested by Oliver (1981), expectations were measured before the shopping experience rather than in retrospect.

As shown in Table 1, online satisfaction was measured using the difference between 8 items of the EtailQ (expectations - perceptions), which corresponds to the two dimensions that could now be judged by the participant: website design and privacy/security. Expectations about the rest of the process (offline) were asked using the 6 other items of the EtailQ in «expectation mode», which corresponds to the following dimensions: fulfilment and customer service.

Online repurchase intention was measured using Limayen et al.'s 3 items scale (2000). Intentions to recommend was measured with two items, adapted from Baumann et al.

(2007)' willingness to recommend scale (I would recommend doing online grocery shopping on this website to my friends, If my friends were looking for a way to shop online, I would tell them to try this website) (see Table 1).

Implicit Measurement

User Experience (UX) was measured with 2 types of physiological signals: automatic facial analysis (valence) and electrodermal activity (arousal). Affective arousal measured the consumers 'emotional sweating' through skin conductance (Boucsein, 2012). Valence was obtained from their emotional state and represented the positivity or negativity of emotions felt by consumers. The automatic facial analysis software FaceReader (Noldus, Wageningen, Netherlands) allowed us to observe the strengths and quantity of emotional reactions (Cohn & De la Torre, 2014; Loijens & Krips, 2018). The online facial expression recognition used by the software recognizes the following universal emotions: happy, angry, sad, surprised, scared, disgust and neutral (Den Uyl & Van Kuilenburg, 2005). At all times, the software analyses facial features locations to determine which of the emotions are felt by the participants (Goldberg, 2012). Emotional valence is therefore a quantitative, linear combination of the output values of facial analysis, with values ranging from positive (+1) to negative (-1) valence (Goldberg, 2012). Simply put, emotional valence is the intensity of the emotion 'happy' minus the intensity of the negative expression with the highest intensity (Loijens & Krips, 2018). Lewinski et al., (2014) validated the FaceReader software results by comparing the matching scores for recognition of facial expressions and the Facial Action Coding System (FACS) index of agreement (Ekman & Friesen, 1978) for both the software and humans, concluding that FaceReader is as good at recognizing emotions as humans.

The skin conductance level, or electrodermal amplitude, was used to measure the degree to which a participant was excited or apathetic. It was obtained through measurement of electrodermal activity linked to electrical conductance of the skin (Boucsein, 2012). As it is well established that sweat gland activity is linked to electrodermal response (Fowles, 1986), using skin conductance has become increasingly popular and is now an established psychological method (Bach, 2016). To measure skin conductance, skin electrodes or usually placed on the palm on the hands as this allows for the measurement of sweat gland activity (Malmivuo, Malmivuo, & Plonsey, 1995). Sweat glands are known to serve thermoregulation (Bach, 2016). However, in some body parts such as the palmar region, sweating increases under particular states, such as affective arousal (Bach, 2016). This particular state of sweating has been termed « emotional sweating » (Boucsein, 2012). More sweat in the palm of the hand means more electrolytes and therefore, lower resistance, allowing to calculate the level of effective arousal based on the conductance of the skin (Bach, 2016). Electrodermal activity was measured with the AcqKnowledge software (BIOPAC, Goleta, USA). A very low constant voltage is applied to the two electrodes placed on participants' palms and results in a current flowing, that is measured and converted to a conductance in accordance with Ohm's law (BIOPAC, 2015). Since BIOPAC is a computer-based system, the analog skin conductance signal is digitized and stored, allowing the researcher to select specific time points (Cacioppo et al., 2007). Finally, Media Recorder (Noldus, Wageningen, Netherlands) was used to record all interactions while Observer XT (Noldus, Wageningen, Netherlands) allowed a synchronisation of the AcqKnowledge, FaceReader, and Media Recorder software.

In order to better evaluate consumers' emotions, particularly negative ones, we chose to build upon a psychophysiological study by Lewinski et al. (2014), which used the top 10% peak values of facial expressions of emotions to perform analysis in order to evaluate emotions felt during advertisement[4]. Therefore, we used the sum of negative emotional responses, by identifying every second where the user was in its ninetieth percentile of EDA (i.e., high arousal), compared to his baseline state and in the tenth percentile of valence (i.e., large negative valence), meaning he felt an intense negative emotion. We chose to look only at the sum of intense negative moments in the shopping task, therefore excluding the tasks such as account creation, payment, and order verification, that are common to most e-commerce websites. The frequencies of intense negative emotions were later calculated using statistical software SAS 9.4.

[4] Analysis were also conducted using the 5% and 15% threshold, showing similar results.

Table 1: Operationalization of Constructs

Construct	Moment of Measurement	Definition	Measure
Negative emotions	Purchase Stage	Frequency of seconds of negative emotions in the shopping task, defined by a combination of low valence (lowest 10%) and high arousal (highest 10%)	• Valence: Facial expression recognition software FaceReaderTM (Noldus, Wageningen, Netherlands) • Arousal: Electrodermal amplitude with the AcqKnowledge software (BIOPAC, Goleta, USA)
Online Satisfaction	Pre and Post Purchase Stage	Users' evaluation toward the **online** process (Perceived performance – expectations)	Items evaluated on a 7-point Likert scale • The website provides in-depth information. • The site doesn't waste my time. • It is quick and easy to complete a transaction at this website. • This website has good selection. • The level of personalization is about right, not too much and not too little.

			- I feel like my privacy is protected at this website. - I feel safe in my transactions with this website. - The website has adequate security features. (Wolfinbarger and Gilly, 2003)
Offline Satisfaction	Post Purchase and Post Fulfilment Stage	Users' evaluation toward the **offline** process (Perceived performance – expectations)	Items evaluated on a 7-point Likert scale - You get what you ordered from this site. - The product(s) is/are delivered by the time promised by the company. - The product(s) that came was/were represented accurately by the website. - The company is willing and ready to respond to customer needs. - When you have a problem, the website shows a sincere interest in solving it. - Inquiries are answered promptly. (Wolfinbarger and Gilly, 2003)

Intentions to repurchase & intention to recommend	Post Fulfilment Stage	Users' intentions at the end of the process	Online repurchase intentions scale (Items evaluated on a 7-point Likert scale) • I anticipate to repurchase from this internet store in the near future. • It is likely that I will repurchase from this internet store in the near future. • I expect to repurchase from this internet store in the near future. (Limayen et al., 2000) Intentions to recommend scale (Items evaluated on a 7-point Likert scale) • I would recommend doing online grocery shopping on this website to my friends. • If my friends were looking for a way to shop online, I would tell them to try this website. Adapted from Baumann et al. (2007)' willingness to recommend scale

Analysis

Sample Statistics

Participants were highly comparable in terms of age and level of education. We didn't see any significant difference between their online shopping habits. To assess potential differences between websites, a repeated measure ANOVA was performed using SUS's 10 items mean evaluated by 10 experts and no significant difference was found. A total of 45 participants were evaluated, but 2 participants were lost between the ordering and fulfilment stage.

Direct Effects

Direct effects have been tested using simple linear regressions. Hierarchical linear regressions were also used to evaluate the proportion of variance explained by each construct.

Mediation Effects

Mediation effects have been tested using Model 4 of the PROCESS macro (Hayes, 2013) in SPSS. 5000 bootstrap samples have been used. We used Preacher and Hayes' method, bootstrapping the sampling distribution of ab in order to derive a confidence interval as well as an estimated standard error, for both 95% and 99% confidence interval (Preacher & Hayes, 2004). This approach is non-parametric and can therefore by applied to small samples with more confidence, which is the case here. Analysis were conducted through bootstrapping with 95% confidence intervals.

Results

Hypothesis testing

Hypotheses 1_a-1_d suggest relationships between both dimensions of online satisfaction and offline satisfaction and intentions to repurchase. Simple linear regressions show that online website design (t=3.30, p=0.002, β=0.338), online security (t=1.77, p=0.042, β=0.148), offline fulfilment (t=3,15, p=0,002, β=0.374) and offline customer service (t=3.65, p=0.001, β=0.439) all positively influence intentions to repurchase. Moreover, a hierarchical linear regression shows online satisfaction accounts for 23% of the total variance in intentions to repurchase in the model while jointly, online and offline satisfaction account for 33%, both variables having a significant impact. Therefore, all four hypotheses are supported.

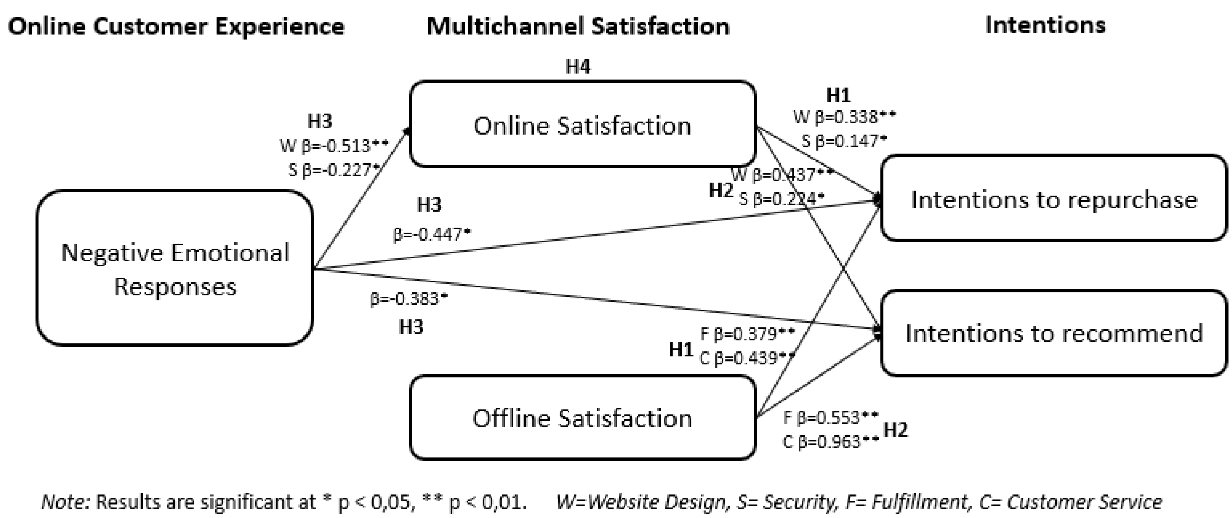


Figure 4: Results

Hypotheses 2_a-2_d suggest relationships between both dimensions of online satisfaction and offline satisfaction and intentions to recommend. Simple linear regressions show that online website design (t=3.41, p=0.001, β=0.437), online security (t=2.17, p=0.018, β=0.224), offline fulfillment (t=3.88, p=0.000, β=0.553) and offline customer service (t=10.97, p=0.000, β=0.963) all positively influence intentions to recommend. Moreover, a hierarchical linear regression shows that online satisfaction accounts for 26% of the total variance in intentions to recommend in the model while jointly, online and offline satisfaction account for 76%, both variables having a significant impact. Therefore, all four hypotheses are supported.

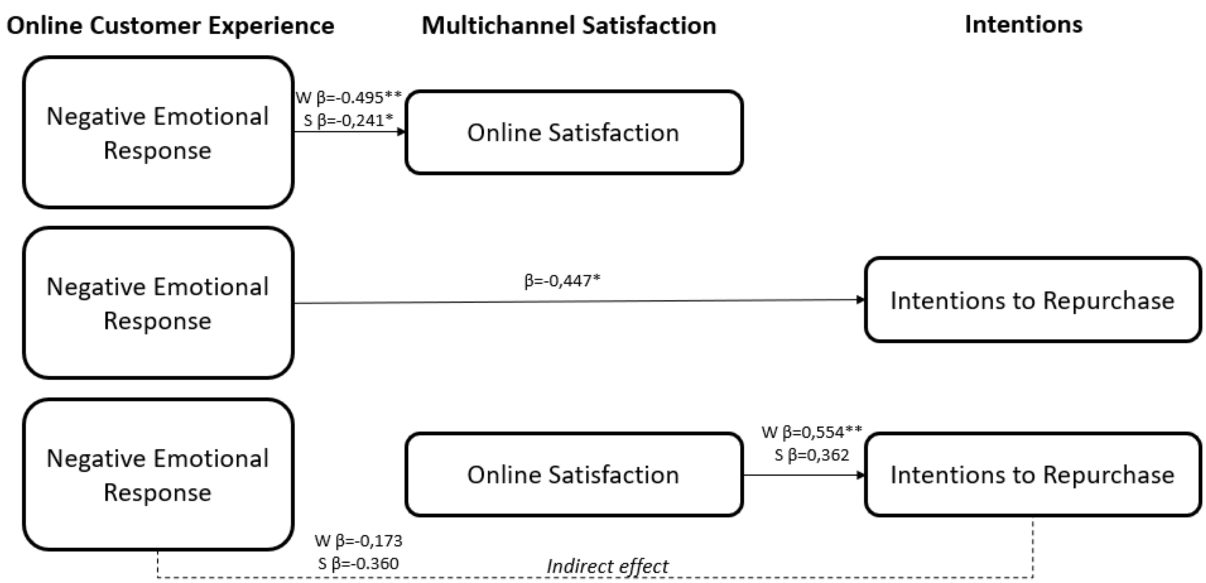
Hypotheses 3_a-1_d examine the effect of negative emotions in the shopping task on online satisfaction as well as on intentions to repurchase and recommend. Simple linear regressions show that negative emotions during the shopping task have a negative impact on online website design (t=-3.36, p=0.001, β=-0.513), online security (t=-1.84, p=0.036, β=-0.227), intentions to repurchase (t=-1.95, p=0.029, β=-0.447) and intentions to recommend (t=-2.12, p=0.020, β=-0.383), supporting all four hypotheses.

Hypotheses 4_a-4_d postulate the mediating effect of both dimensions of online satisfaction between consumers' negative emotions in the shopping task and intentions to recommend and repurchase. Mediation analysis were conducted using both online satisfaction toward website design and security as mediators, as well as both intentions to repurchase and intentions to recommend as the dependent variables, in 4 separate mediation analysis.

In the first mediation analysis, shown in Figure 5, the negative emotions have a negative influence on online satisfaction toward website design (t=-3.10, p=0.002, β=-0.495) and intentions to repurchase (t=-1,94, p=0.029, β=-0,447), when tested separately. However, when tested together, the negative emotions have no effect on intentions to repurchase (t=-0,72, p=0.237, β=-0.173), but online satisfaction toward the website design positively influenced intentions to repurchase (t=2.63, p=0.006, β=0.554). With a 95% bootstrap confidence interval, the interval estimate is -0.645 to -0.035, showing a significant indirect effect. Therefore, we conclude that online satisfaction toward the website design was a full mediator in the relationship between the negative emotions and intentions to repurchase, supporting H4a.

In the second mediation analysis, shown in Figure 6, negative emotions affect online satisfaction toward security (t=-1.88, p=0.033, β=-0.241) and intentions to repurchase (t=-1.95, p=0.029, β=-0.447). As online satisfaction toward security was entered in the model, both variables became nonsignificant (online satisfaction toward security → intentions to repurchase: t=1.29, p=0.101, β=0.362; negative emotions → intentions to repurchase: t=-1.51, p=0.069, β=-0.360). With a 95% bootstrap confidence interval, the interval estimate is -0.295 to 0.056, showing a non-significant indirect effect. Therefore, we conclude that there is no mediation effect of satisfaction toward online security in the relationship between the negative emotions and intentions to repurchase and that H4b is not supported.

In the third mediation analysis, also shown in Figure 5, the negative emotions have a negative influence on online satisfaction toward website design (t=-3.10, p=0.002, β=-0,495), and intentions to recommend (t=-2.11, p=0.020,β=-0.383), when tested separately. However, when tested together, the negative emotions have no effect on intentions to recommend (t=-0.88, p=0.193, β=-0.164), but online satisfaction toward the website design positively influence intentions to recommend (t=2.68, p=0.005, β=0.442). With a 95% bootstrap confidence interval, the interval estimate is -0.452 to -0.027, showing a significant indirect effect. Therefore, we concluded that online satisfaction toward the website design is a full mediator in the relationship between the negative emotions and intentions to recommend, supporting H4c.

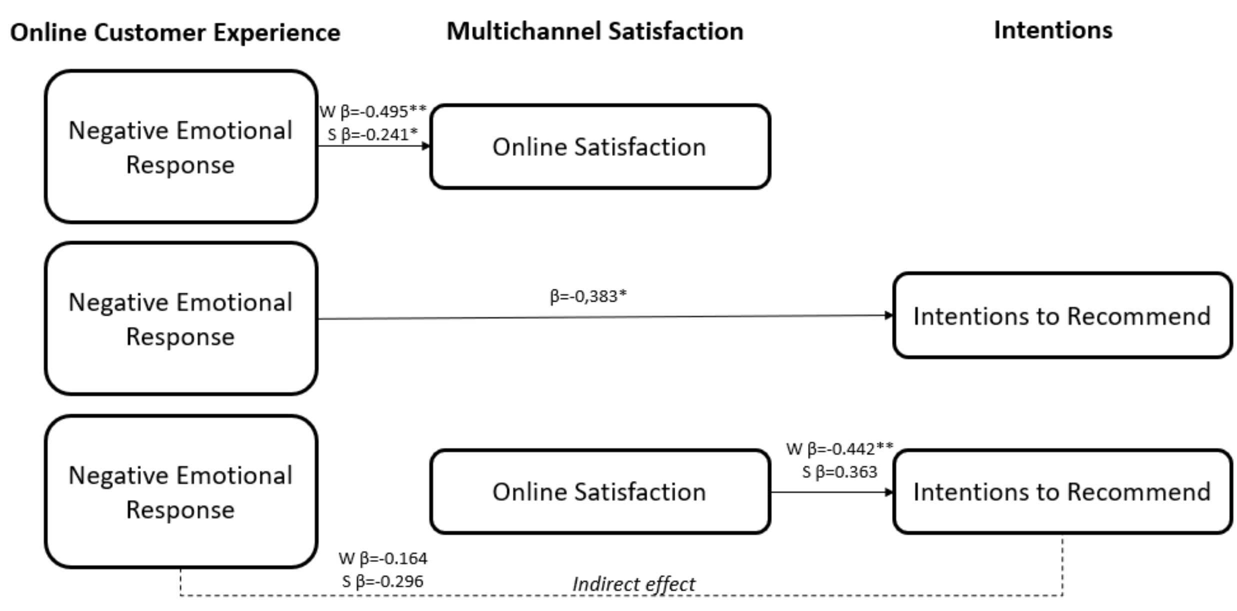


Note: Results are significant at * p < 0,05, ** p < 0,01. *W=Website Design, S= Security*

Figure 5: Mediation Analysis for Intentions to Repurchase

In the final mediation analysis, shown in Figure 6, negative emotions affect online satisfaction toward security (t=-1.88, p=0.033, β=-0.241) and intentions to recommend (t=-2.11, p=0.020, β=-0.383). As online satisfaction toward security was entered in the model, both variables became nonsignificant (online satisfaction toward security → intentions to recommend: t=1.67, p=0.051, β=0.363; negative emotions → intentions to recommend: t=-1.60, p=0.059, β=-0.296). With a 95% bootstrap confidence interval, the interval estimate is -0.310 to 0.047, showing a non-significant indirect effect. Therefore, we conclude that there is no mediation effect of satisfaction toward online security in the relationship between the negative emotions and intentions to recommend and that H4d is not supported.

To summarize, results show the indirect effect of the negative emotions in the shopping task on intentions to recommend and repurchase mediated through online website design (H4a and H4c). However, online security was not found to be a mediator (H4b and H4d), supporting two out of the four hypotheses. Results for mediation effects are summarized in Figures 5 and 6.



Note: Results are significant at * p < 0,05, ** p < 0,01. *W=Website Design, S= Security*

Figure 6: Mediation Analysis for Intentions to Recommend

Discussion

Results of the study suggest positive relationships between both online satisfaction and offline satisfaction and intentions to repurchase (H1a-d), as well as between both online satisfaction and offline satisfaction and intentions to recommend (H2a-d). Furthermore, results highlight the effect of negative emotions in the shopping task on online satisfaction, as well as on intentions to repurchase and recommend (H3a-d). Finally, results suggest a mediating effect of one of the two dimensions of online satisfaction, satisfaction toward the website design, on the relationship between consumers' negative emotions during the shopping task and intentions to recommend and repurchase (H4a&c). However, the mediating effect of the second dimension of online satisfaction, satisfaction toward security, is not supported (H4b&d).

This study has theoretical implications. First, by investigating the whole multichannel process into a single consumer journey, results provide a comprehensive understanding of the formation of customer satisfaction, from pre-purchase expectations to post-purchase fulfilment, which was lacking in current literature (D. J. Kim et al., 2009). Accordingly, using the expectation-confirmation theory combined with the etailQ scale allowed us to gain a better understanding of the role of each dimension in the formation of satisfaction. In online satisfaction, the website design was found to influence both intentions to repurchase and intentions to recommend than security. Using the expectation confirmation theory (Oliver, 1980) allowed us to see that security might not be a strong predictor because of its low variance, as all customers now expect the highest level of security from most e-commerce websites. In offline satisfaction, customer service was also found to be more correlated with both intentions to repurchase and intentions to recommend than fulfillment, which suggest that even though people do an online transaction, a high importance is still given to human interactions. The results also show that overall, offline satisfaction explains better intentions to recommend than online satisfaction, which is coherent with previous research suggesting that final satisfaction, or the last mile, is a main determinant of intentions (Fernie & Sparks, 2018). However, regarding intentions to repurchase, online satisfaction explains a higher proportion of variance, which goes against some of Oliver's previous work (1993), showing that final satisfaction is a main determinant of whether consumers will repurchase or not.

Second, results suggest that psychophysiological negative emotions influence satisfaction toward the online experience, but also influences consumers' intentions to repurchase and recommend. As users' satisfaction and intentions are crucial in marketing research (Kumar & Oliver, 1997; Richard L. Oliver, 1993), integrating implicit measures to well-established antecedents to users' satisfaction and intentions allows a better understanding of consumers' evaluation by seeing how implicit and explicit measures interact in the formation of satisfaction and intentions (de Guinea et al., 2014). Results suggest that negative emotions, assessed by psychophysiological measures, in the shopping task not only negatively influence satisfaction toward the website, but also consumers' intentions to repurchase and recommend the website. This is coherent with prior research showing that negative emotions affect satisfaction (Dube-Rioux, 1990; Hui & Tse, 1996; Richard L. Oliver, 1993; Price et al., 1995) as well as future intentions (Inman et al., 1997; Zeelenberg & Pieters, 2004). Furthermore, we found that the marginal effect of an additional intense negative moment in the shopping task significantly impacted all constructs. Each additional intense negative moment decreased the total perceived score for the following constructs: satisfaction toward website design (-0,298), satisfaction toward online security (-0,132), intentions to repurchase (-0,260) and intentions to recommend (-0,223). The scores are on 7-point Likert scales. This implies that for satisfaction toward the website design, for example, each negative moment decreases the satisfaction score, whose values range is between 1 and 7, by 0,290. Although no previous research has been done studying the effects of the frequency of intense negative emotions, results are coherent with work by Baumeister et al. (2001) that show that bad memories are more easily remembered than good ones as well as work by Ariely (1998), suggested that the intensity of the emotions felt play a role in its retrospective evaluation of a moment.

Third, results show that satisfaction mediates the relationship between the psychophysiological emotional responses and intentions, which is coherent with existing literature (Gustafsson et al., 2005; Hsieh et al., 2012). More precisely, results show the indirect effect of negative emotions during the shopping task on intentions to recommend and repurchase mediated through online website design. Those results show that both explicit and implicit measures are important in predicting satisfaction and intentions and provides a deeper understanding of the relationship between both types of measures, which is coherent with recent research showing that implicit responses, often unconscious, can interact with explicit ones in the formation of beliefs (de Guinea et al., 2014). Results also highlight the mediating role of online satisfaction on intentions, suggesting that consumers' perceptions can mediate the effects of unconscious processes of the experience.

This study also has managerial implications. The ecological validity of this study was very high, as consumers were asked to shop according to their needs, using their own credit card and picking up their order from the store, therefore engaging in the full process from beginning to end.

First, this study shows the importance of the website design dimension to decrease the occurrence of negative emotions in the shopping task. Many executives have been shown to focus on attracting customers on their website rather than retaining them (Reichheld & Schefter, 2000), which often leads to customers leaving the website, unsatisfied with their experience. The results of this study allow retailers, and more specifically, grocers, to devote their valuable resources to the attributes that really make a difference, such as the website design, in order to gain customer satisfaction (G.-G. Lee & Lin, 2005).

Second, it also shows the trickle-down effect of online negative emotions on satisfaction and intentions to repurchase and recommend. While the relationship between satisfaction and higher profit has been established many times (E. W. Anderson & Mittal, 2000; Storbacka, Strandvik, & Grönroos, 1994), intentions can be as important. Returning customers can be five times more profitable than new ones (Gupta & Kim, 2007). Intentions to repurchase are also very profitable: referrals have been shown to be more effective in e-commerce than in traditional commerce, « word of mouse » spreading much faster than word of mouth (Reichheld & Schefter, 2000). Since results show that online satisfaction is critical, both short term and long term, investing in website design and security is crucial.

Third, this study points out the importance of offline satisfaction on intentions. As offline (dis) satisfaction is achieved at the very end, after the consumer picks up his order from the store, this shows the importance of the last mile in consumer repurchase and recommend intentions. Studying the last mile fulfilment in store pick-up method, allows grocers to pinpoint the most important steps to improve and optimize the process (Hübner, Kuhn, & Wollenburg, 2016).

Finally, this study highlights the importance of the omnichannel experience and their integration, so that customers get a seamless journey while having both online and physical experiences in the same transaction process. Research shows that « brick and click» businesses have many competitive advantages compared to pure-play businesses, such as offering an extended selection online, flexibility, convenience and trustworthiness of a known brand (Min & Wolfinbarger, 2005). As online retailers, like Amazon, already have high market share (Kahn et al., 2018), this study shows the need for retailers to use their competitive advantages. This study shows the importance for multichannel grocers to develop appropriate strategies to provide high service qualities on all channels.

Limitations and Future Research

Limitations need to be acknowledged. First, the limited number of participants as well as the convenient sampling limits the generalization of the results. The number of participants was restrained by physiological data collection complexity as well as the multichannel aspect of the study, that had to be done over a certain period of time. Moreover, this study recruited participants with our institution research panel, therefore recruiting participants from the same institution. Thus, future research could replicate this study with a larger sample size and with a more diversified sample, by studying different demographics. Additionally, participants were forced to complete a purchase on one of three randomly assigned websites. This decreases the ecological validity, as the grocer was not necessarily the consumers' first choice and other factors, such as proximity to home, could have influenced his satisfaction. Therefore, future research could control for those potential biases. Finally, this study was done on utilitarian websites and involved long transactions, which is not the case for all e-commerce transactions. As the frequency of intense negative emotional responses seem to be correlated with many different constructs, future research is needed to improve the generalizability of the research model. Future research in hedonic contexts would help gain a better understanding of consumers' psychophysiological states. Furthermore, as negative emotions sometimes result from tiredness rather than a problem with the interaction, future research should examine negative emotions in shorter transactions to see if results remain similar.

References

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, 36(7), 623-636.

Albert, W., & Tullis, T. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics:* Newnes.

Anderson, E. W., & Mittal, V. (2000). Strengthening the satisfaction-profit chain. *Journal of Service research*, 3(2), 107-120.

Anderson, E. W., & Salisbury, L. C. (2003). The formation of market-level expectations and its covariates. *Journal of Consumer Research*, 30(1), 115-124.

Anderson, E. W., & Sullivan, M. W. (1993). The antecedents and consequences of customer satisfaction for firms. *Marketing science*, 12(2), 125-143.

Anderson, R. E., & Srinivasan, S. S. (2003). E-satisfaction and e-loyalty: A contingency framework. *Psychology and Marketing*, 20(2), 123-138. doi:10.1002/mar.10063

Anesbury, Z., Nenycz-Thiel, M., Dawes, J., & Kennedy, R. (2016). How do shoppers behave online? An observational study of online grocery shopping. *Journal of Consumer Behaviour*, 15(3), 261-270.

Ariely, D. (1998). Combining experiences over time: the effects of duration, intensity changes and on-line measurements on retrospective pain evaluations. *Journal of Behavioral Decision Making*, 11(1), 19-45.

Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. A*cademy of Management Journal*, 24(2), 377-385.

Bach, D. R. (2016). Skin conductance measures in neuroeconomic research Neuroeconomics (pp. 345-357): Springer.

Baumann, C., Burton, S., Elliott, G., & Kehr, H. M. (2007). Prediction of attitude and behavioural intentions in retail banking. *International journal of bank marketing*, 25(2), 102-116.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, 5(4), 323-370.

Bhattacherjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *Mis Quarterly*, 351-370.

Bigné, J. E., Andreu, L., & Gnoth, J. (2005). The theme park experience: An analysis of pleasure, arousal and satisfaction. *Tourism management*, 26(6), 833-844.

Bloomberg (2019). *Online food and beverage market share in the United States in 2016, by company*. In Statista - The Statistics Portal.

Boucsein, W. (2012). *Electrodermal activity*: Springer Science & Business Media.

Boulding, W., Kalra, A., Staelin, R., & Zeithaml, V. A. (1993). A dynamic process model of service quality: from expectations to behavioral intentions. *Journal of marketing research*, 30(1), 7-27.

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.

Brady, M. K., Voorhees, C. M., Cronin Jr, J. J., & Bourdeau, B. L. (2006). The good guys don't always win: the effect of valence on service perceptions and consequences. *Journal of Services Marketing*, 20(2), 83-91.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4-7.

Cai, Shaohan et Jun, Minjoon. (2003). Internet users' perceptions of online service quality: a comparison of online buyers and information searchers. *Managing Service Quality: An International Journal*, 13(6), 504-519.

Cao, Y., Gruca, T. S., & Klemz, B. R. (2003). Internet pricing, price satisfaction, and customer satisfaction. *International Journal of Electronic Commerce*, 8(2), 31-50.

Caruana, A., & Ewing, M. T. (2006). The psychometric properties of eTail quality: An international investigation across product categories. *International Marketing Review*, 23(4), 353-370.

Casaló, L. V., Flavián, C., & Guinalíu, M. (2008). The role of satisfaction and website usability in developing customer loyalty and positive word-of-mouth in the e-banking services. *International journal of bank marketing*, 26(6), 399-417.

Chea, S., & Luo, M. M. (2008). Post-adoption behaviors of e-service customers: The interplay of cognition and emotion. *International Journal of Electronic Commerce*, 12(3), 29-56.

Chen, Qimei, & Wells, William D. (1999). Attitude toward the site. *Journal of Advertising Research*, 39(September/October), 27–37.

Chiu, C.-M., Hsu, M.-H., Sun, S.-Y., Lin, T.-C., & Sun, P.-C. (2005). Usability, quality, value and e-learning continuance decisions. *Computers & education*, 45(4), 399-416.

Citrin, A. V., Stem Jr, D. E., Spangenberg, E. R., & Clark, M. J. (2003). Consumer need for tactile input: An internet retailing challenge. *Journal of Business Research*, 56(11), 915-922.

Cockburn, A., Quinn, P., & Gutwin, C. (2015). *Examining the peak-end effects of subjective experience*. Paper presented at the Proceedings of the 33rd annual ACM conference on human factors in computing systems.

Cockburn, A., Quinn, P., & Gutwin, C. (2017). The effects of interaction sequencing on user experience and preference. *International Journal of Human-Computer Studies*, 108, 89-104.

Cohn, J. F., & De la Torre, F. (2014). Automated face analysis for affective *The Oxford handbook of affective computing* (pp. 131).

Courtemanche, F., Léger, P.-M., Dufresne, A., Fredette, M., Labonté-LeMoyne, É., & Sénécal, S. (2017). Physiological heatmaps: a tool for visualizing users' emotional reactions. *Multimedia Tools and Applications*, 77(9), 11547-11574. doi:10.1007/s11042-017-5091-1

Cristobal, E., Flavián, C., & Guinaliu, M. (2007). Perceived e-service quality (PeSQ) Measurement validation and effects on consumer satisfaction and web site loyalty. *Managing Service Quality: An International Journal*, 17(3), 317-340.

Cronin Jr, J. J., & Taylor, S. A. (1994). SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality. *Journal of marketing*, 58(1), 125-131.

Dabholkar, P. A., Shepherd, C. D., & Thorpe, D. I. (2000). A comprehensive framework for service quality: an investigation of critical conceptual and measurement issues through a longitudinal study. *Journal of Retailing*, 76(2), 139-173.

De Guinea, A. O., Titah, R., & Léger, P.-M. (2013). Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research. *Computers in Human Behavior*, 29(3), 833-844.

de Guinea, A. O., Titah, R., & Léger, P.-M. (2014). Explicit and implicit antecedents of users' behavioral beliefs in information systems: A neuropsychological investigation. *Journal of Management Information Systems*, 30(4), 179-210.

Delone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, 19(4), 9-30.

Den Uyl, M., & Van Kuilenburg, H. (2005). *The FaceReader: Online facial expression recognition*. Paper presented at the Proceedings of measuring behavior.

Desrochers, C.; Léger, P.-M.; Mirhoseini, S.; Fredette, M.; Sénécal, S.(Forthcoming). The Arithmetic Complexity of Online Grocery Shopping: The Moderating Role of Product Pictures. *Industrial Management & Data Systems*.

Dirican, A. C., & Göktürk, M. (2011). Psychophysiological measures of human cognitive states applied in human computer interaction. *Procedia Computer Science*, 3, 1361-1367.

Dube-Rioux, L. (1990). The power of affective reports in predicting satisfaction judgments. *ACR North American Advances.*

Ekman, P., & Friesen, W. V. (1978). *Manual for the facial action coding system*: Consulting Psychologists Press.

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency *Advances in experimental social psychology* (Vol. 14, pp. 161-202): Elsevier.

Fernie, J., & Sparks, L. (2018). *Logistics and retail management: emerging issues and new challenges in the retail supply chain*: Kogan page publishers.

Festinger, L. (1957). *A theory of cognitive dissonance* (Vol. 2): Stanford university press.

Forrester. (2018). *Online Grocery Retail Forecast, 2017 to 2022 (Global),* Forrester Report, Cambridge, MA.

Fowles, D. C. (1986). The eccrine system and electrodermal activity. Psychophysiology: *Systems, processes, and applications*, 1, 51-96.

Goldberg, J. H. (2012). *Relating perceived web page complexity to emotional valence and eye movement metrics*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Gupta, S., & Kim, H.-W. (2007). The moderating effect of transaction experience on the decision calculus in on-line repurchase. *International Journal of Electronic Commerce*, 12(1), 127-158.

Gustafsson, A., Johnson, M. D., & Roos, I. (2005). The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. *Journal of marketing*, 69(4), 210-218.

Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research*, 62(5), 565-571.

Han, H., Back, K.-J., & Barrett, B. (2009). Influencing factors on restaurant customers' revisit intention: The roles of emotions and switching barriers. *International Journal of Hospitality Management*, 28(4), 563-572.

Hartson, R., & Pyla, P. S. (2012). *The UX Book: Process and guidelines for ensuring a quality user experience:* Elsevier.

Hassenzahl, M. (2013). User experience and experience design. *The encyclopedia of human-computer interaction,* 2.

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität *Mensch & computer* 2003 (pp. 187-196): Springer.

Hsieh, Y.-C., Roan, J., Pant, A., Hsieh, J.-K., Chen, W.-Y., Lee, M., & Chiu, H.-C. (2012). All for one but does one strategy work for all? Building consumer loyalty in multi-channel distribution. *Managing Service Quality: An International Journal,* 22(3), 310-335.

Huang, Y., & Oppewal, H. (2006). Why consumers hesitate to shop online: An experimental choice analysis of grocery shopping and the role of delivery fees. *International journal of Retail & Distribution Management*, 34(4/5), 334-353.

Hübner, A., Kuhn, H., & Wollenburg, J. (2016). Last mile fulfilment and distribution in omni-channel grocery retailing: a strategic planning framework. *International journal of Retail & Distribution Management*, 44(3), 228-247.

Hui, M. K., & Tse, D. K. (1996). What to tell consumers in waits of different lengths: An integrative model of service evaluation. *Journal of marketing*, 60(2), 81-90.

Inman, J. J., Dyer, J. S., & Jia, J. (1997). A generalized utility model of disappointment and regret effects on post-choice valuation. *Marketing science*, 16(2), 97-111.

ISO: ISO 9241-210:2010 (2010). Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems (formerly known as 13407). *International Organization for Standardization* (ISO), Switzerland.

Kahn, B. E., Inman, J. J., & Verhoef, P. C. (2018). Introduction to Special Issue: Consumer Response to the Evolving Retailing Landscape. *Journal of the Association for Consumer Research*, 3(3), 255-259.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological science*, 4(6), 401-405.

Khalifa, M., & Liu, V. (2002). Satisfaction with internet-based services: the role of expectations and desires. *International Journal of Electronic Commerce*, 7(2), 31-49.

Kim, D. J. (2012). An investigation of the effect of online consumer trust on expectation, satisfaction, and post-expectation. *Information Systems and E-Business Management*, 10(2), 219-240.

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2009). Trust and Satisfaction, Two Stepping Stones for Successful E-Commerce Relationships: A Longitudinal Exploration. *Information Systems Research*, 20(2), 237-257. doi:10.1287/isre.1080.0188

Kim, J., Jin, B., & Swinney, J. L. (2009). The role of etail quality, e-satisfaction and e-trust in online loyalty development process. *Journal of Retailing and Consumer Services*, 16(4), 239-247.

Kumar, A., & Oliver, R. L. (1997). Special Session Summary Cognitive Appraisals, Consumer Emotions, and Consumer Response. *ACR North American Advances*.

Ladhari, R., Souiden, N., & Dufour, B. (2017). The role of emotions in utilitarian service settings: The effects of emotional satisfaction on product perception and behavioral intentions. *Journal of Retailing and Consumer Services*, 34, 10-18.

Langer, T., Sarin, R., & Weber, M. (2005). The retrospective evaluation of payment sequences: duration neglect and peak-and-end effects. *Journal of Economic Behavior & Organization*, 58(1), 157-175.

Lee, G.-G., & Lin, H.-F. (2005). Customer perceptions of e-service quality in online shopping. *International journal of Retail & Distribution Management*, 33(2), 161-176.

Lee, H. S. (2013). Major moderators influencing the relationships of service quality, customer satisfaction and customer loyalty. *Asian Social Science*, 9(2), 1.

Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, 7(4), 227.

Lewinski, P., Fransen, M. L., & Tan, E. S. (2014). Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli. *Journal of Neuroscience, Psychology, and Economics*, 7(1), 1.

Li, H., Aham-Anyanwu, N., Tevrizci, C., & Luo, X. (2015). The interplay between value and service quality experience: e-loyalty development process through the eTailQ scale and value perception. *Electronic Commerce* Research, 15(4), 585-615.

Liao, C., Palvia, P., & Lin, H.-N. (2006). The roles of habit and web site quality in e-commerce. *International Journal of Information Management*, 26(6), 469-483. doi:10.1016/j.ijinfomgt.2006.09.001

Lien, C.-H., Wen, M.-J., & Wu, C.-C. (2011). Investigating the Relationships among E-Service Quality, Perceived Value, Satisfaction and Behavioral Intentions in Taiwanese Online Shopping. *Asis Pacific Management Review*, 16(3), 211-223.

Liljander, V., & Strandvik, T. (1997). Emotions in service satisfaction. *International journal of service industry management*, 8(2), 148-169.

Lin, C., & Lekhawipat, W. (2014). Factors affecting online repurchase intention. *Industrial Management & Data Systems*, 114(4), 597-611. doi:10.1108/imds-10-2013-0432

Lin, C., & Lekhawipat, W. (2016). How customer expectations become adjusted after purchase. *International Journal of Electronic Commerce,* 20(4), 443-469.

Loiacono, E. T., Watson, R. T., & Goodhue, D. L. (2002). WebQual: A measure of website quality. *Marketing theory and applications*, 13(3), 432-438.

Loijens, L., & Krips, O. (2018). FaceReader Methodology Note. *A white paper by Noldus Information Technology.*

Lourties, S.; Léger, P. M.; Sénécal, S.; Fredette, M.; & Chen, S. L. (2018). *Testing the Convergent Validity of Continuous Self-Perceived Measurement Systems: An Exploratory Study*. In International Conference on HCI in Business, Government, and Organizations , 132-144.

Malmivuo, P., Malmivuo, J., & Plonsey, R. (1995*). Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*: Oxford University Press, USA.

Medina, C., Rufín, R., & Rey, M. (2015). Mediating relationships in and satisfaction with online technologies: communications or features beyond expectations? *Service Business*, 9(4), 587-609.

Meyers-Levy, J., & Malaviya, P. (1999). Consumers' processing of persuasive advertisements: An integrative framework of persuasion theories. *Journal of marketing*, 63(4_suppl1), 45-60.

Min, S., & Wolfinbarger, M. (2005). Market share, profit margin, and marketing efficiency of early movers, bricks and clicks, and specialists in e-commerce. *Journal of Business Research*, 58(8), 1030-1039.

Mooradian, T. A., & Olver, J. M. (1997). "I can't get no satisfaction:" The impact of personality and emotion on postpurchase processes. *Psychology & marketing*, 14(4), 379-393.

Murfield, M., Boone, C. A., Rutner, P., & Thomas, R. (2017). Investigating logistics service quality in omni-channel retailing. *International Journal of Physical Distribution & Logistics Management*, 47(4), 263-296.

Myers, J. H. (1991). Measuring customer satisfaction: is meeting expectations enough? *Marketing Research*, 3(4).

Nielson. (2017). *What's in-store for Online Grocery Shopping, Omnichannel Strategies to Reach Crossover Shoppers*. Nielson Global Connected Commerce Report, New York.

O. Pappas, I., G. Pateli, A., N. Giannakos, M., & Chrissikopoulos, V. (2014). Moderating effects of online shopping experience on customer satisfaction and repurchase intentions. *International journal of Retail & Distribution Management*, 42(3), 187-204.

Oliver, R. L. (1980). A Cognitive Model for the Antecedents and Consequences of Satisfaction, *Journal of Marketing Research,* (17), 460-469.

Oliver, R. L. (1981). Measurement and evaluation of satisfaction processes in retail settings. *Journal of Retailing.*

Oliver, R. L. (1993). Cognitive, Affective, and Attributes Bases of the Satisfaction Response. *Journal of Consumer Resear*ch, 20(3), 418-430.

Oliver, R. L., & Winer, R. S. (1987). A framework for the formation and structure of consumer expectations: Review and propositions. *Journal of economic psychology*, 8(4), 469-499.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of marketing*, 49(4), 41-50.

Pew Research Center. (2018). *Adult internet usage penetration in the United States from 2000 to 2018, by age group*. In Statista - The Statistics Portal.

Phillips, D. M., & Baumgartner, H. (2002). The role of consumption emotions in the satisfaction response. *Journal of Consumer Psychology*, 12(3), 243-252.

Piotrowicz, W., & Cuthbertson, R. (2014). Introduction to the special issue information technology in retail: Toward omnichannel retailing. *International Journal of Electronic Commerce*, 18(4), 5-16.

Poels, K., & Dewitte, S. (2006). How to capture the heart? Reviewing 20 years of emotion measurement in advertising. *Journal of Advertising Research*, 46(1), 18-37.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(3), 715-734.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers*, 36(4), 717-731.

Price, L. L., Arnould, E. J., & Tierney, P. (1995). Going to extremes: Managing service encounters and assessing provider performance. *Journal of marketing*, 59(2), 83-97.

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2001). *The retina.* Neuroscience. 5th ed. Sunderland, MA: Sinauer Associates.

Reichheld, F. F., & Schefter, P. (2000). E-loyalty: your secret weapon on the web. *Harvard business review*, 78(4), 105-113.

Riedl, R., Fischer, T., & Léger, P.-M. (2017). A decade of NeuroIS research: status quo, challenges, and future directions.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, 110(1), 145.

Rust, R. T., & Lemon, K. N. (2001). E-service and the consumer. *International Journal of Electronic Commerce*, 5(3), 85-101.

Rychalski, A., & Hudson, S. (2017). Asymmetric effects of customer emotions on satisfaction and loyalty in a utilitarian service context. *Journal of Business Research*, 71, 84-91.

Santos, J. (2003). E-service quality: a model of virtual service quality dimensions. *Managing Service Quality: An International Journal*, 13(3), 233-246.

Schooler, J. W., & Eich, E. (2000). *Memory for emotional events*.

Seth, N., Deshmukh, S., & Vrat, P. (2005). Service quality models: a review. *International journal of quality & reliability management*, 22(9), 913-949.

Silvestre, A.-L., Sue, V. M., & Allen, J. Y. (2009). If you build it, will they come? The Kaiser Permanente model of online health care. *Health Affairs*, 28(2), 334-344.

Statista. (2017*), U.S. consumers: Online Grocery Shopping-Statistics & Facts*, Hambourg.

Statista (2018*). eCommerce Report 2019* – Electronics & Media, Statista Digital Market Outlook – Segment Report

Storbacka, K., Strandvik, T., & Grönroos, C. (1994). Managing customer relationships for profit: the dynamics of relationship quality. *International journal of service industry management*, 5(5), 21-38.

Taylor, S., & Todd, P. (1995). Decomposition and crossover effects in the theory of planned behavior: a study of consumer adoption intentions. *International Journal of Research in Marketing*, 12, 137-155.

The Hartman Group & FMI (2018). *Share of shoppers who have used an online-only retailer for groceries in the United States in 2018*, by generation. In Statista - The Statistics Portal.

Venkatesh, V., Thong, J. Y., 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.

Walker, W. R., Vogl, R. J., & Thompson, C. P. (1997). Autobiographical memory: Unpleasantness fades faster than pleasantness over time. Applied Cognitive Psychology: *The Official Journal of the Society for Applied Research in Memory and Cognition*, 11(5), 399-413.

Wimmer, B., Wöckl, B., Leitner, M., & Tscheligi, M. (2010). *Measuring the dynamics of user experience in short interaction sequences*. Paper presented at the Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries.

Wirtz, J., Mattila, A. S., & Tan, R. L. (2000). The moderating role of target-arousal on the impact of affect on satisfaction—an examination in the context of service experiences. *Journal of Retailing*, 76(3), 347-365.

Wolfinbarger, M., & Gilly, M. C. (2003). eTailQ: dimensionalizing, measuring and predicting etail quality. *Journal of Retailing*, 79(3), 183-198.

Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services. *International journal of service industry management*, 15(3), 302-326.

Yang, Z., & Peterson, R. T. (2004). Customer perceived value, satisfaction, and loyalty: The role of switching costs. *Psychology & marketing*, 21(10), 799-822.

Yoo, B., & Donthu, N. (2001). Developing a Scale to Measure the Perceived Quality of An Internet Shopping Site (SITEQUAL). *Quarterly Journal of Electronic Commerce*, 2(1), 31-47.

Zaman, B., & Shrimpton-Smith, T. (2006). *The FaceReader: Measuring instant fun of use*. Paper presented at the Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles.

Zboja, J. J., & Voorhees, C. M. (2006). The impact of brand trust and satisfaction on retailer repurchase intentions. *Journal of Services Marketing*, 20(6), 381-390.

Zeelenberg, M., & Pieters, R. (2004). Beyond valence in customer dissatisfaction: A review and new findings on behavioral responses to regret and disappointment in failed services. *Journal of Business Research*, 57(4), 445-455.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through web sites: a critical review of extant knowledge. *Journal of the Academy of Marketing Science*, 30(4), 362-375.

Conclusion

Ce mémoire visait deux objectifs et a été réalisé en deux temps. Le premier objectif était d'introduire et de tester une nouvelle méthode de mesure implicite de l'expérience client en ligne. Le second était de combiner les résultats de celle-ci avec des mesures explicites, afin de tenter de mieux expliquer la satisfaction dans un contexte d'épicerie en ligne. Plus précisément, ce mémoire a permis de comparer l'expérience réellement vécue par les consommateurs, mesurée à l'aide de points de friction psychophysiologiques, avec celle perçue par ceux-ci, mesurée à l'aide d'une entrevue et de questionnaires. Il offre également un complément d'information en présentant pas à pas les étapes à effectuer pour être capable de reproduire la méthode proposée. Finalement, ce mémoire a permis d'avoir une meilleure compréhension du rôle de la satisfaction dans la relation entre l'expérience utilisateur du consommateur et ses intentions ainsi que de l'impact d'une expérience vécue négative sur la satisfaction des consommateurs et leurs intentions.

Une expérience en laboratoire a été réalisée à l'hiver 2018, avec 21 participants, et a été reproduite identiquement à l'hiver 2019, avec 24 participants supplémentaires. Ceux-ci devaient effectuer leur épicerie en ligne sur un des trois sites internet d'épicerie en ligne, auxquels ils étaient assignés aléatoirement, en utilisant leur propre carte de crédit et en magasinant selon leurs besoins réels. Les participants devaient par la suite récupérer leur commande à l'épicerie de leur choix et compléter des questionnaires tout au long du processus (avant l'expérience, après l'expérience, après la collecte de la commande à l'épicerie). Les participants ont été rémunérés avec 60$ en argent comptant, pour compenser pour l'épicerie achetée, ainsi que 10$ en carte cadeau COOP HEC suite à la complétion du dernier formulaire. L'expérience en ligne a été mesurée à l'aide d'un oculomètre, ainsi que du logiciel d'analyse des émotions faciales FaceReader, pour mesurer la valence émotionnelle, et BIOPAC, pour mesurer l'intensité émotionnelle. Par la suite, une courte entrevue a été effectuée afin de comparer l'expérience vécue à l'expérience perçue. Des données par questionnaires ont été collectées à différents moments, afin de mesurer l'évolution de la satisfaction au travers plusieurs canaux, soit en épicerie et en ligne. La première collecte de données a permis la rédaction des articles présentés dans les chapitres 2 et 3, alors que la combinaison de la première et de la deuxième collecte a permis la rédaction de l'article présenté au chapitre 4.

Le présent chapitre rappelle les questions de recherche sur lesquelles se base ce mémoire et présente les principaux résultats obtenus au travers des trois articles. Finalement, les contributions théoriques et les implications managériales sont abordées, ainsi que les limites de l'étude et les futures avenues de recherche.

Rappel des questions de recherche et principaux résultats

Les résultats de l'étude nous ont permis de fournir des réponses aux questions de recherche de ce mémoire. Celles-ci ont été répondues en deux temps. Dans un premier temps, ce mémoire avait pour but de répondre aux questions suivantes:

Q1: Comment les mesures psychophysiologiques peuvent-elles nous permettre d'identifier les points de friction implicites dans le parcours client en ligne?

Q2: Les consommateurs sont-ils en mesure d'identifier rétrospectivement avec précision les points de friction de leur propre parcours en ligne?

Celles-ci ont été répondues dans les chapitres 2 et 3 de ce mémoire. Les résultats démontrent qu'en combinant la valence émotionnelle et l'intensité émotionnelle, il est possible d'identifier précisément l'occurrence de points de friction psychophysiologiques implicites dans le parcours du consommateur. De plus, la méthode utilisée, présentée étape par étape dans le chapitre 3, permet de représenter visuellement les résultats à l'aide d'une carte du parcours du consommateur, présentant l'évolution de la valence émotionnelle (axe des y) et de l'intensité émotionnelle (grosseurs des points) sur le temps (axe des x). En identifiant les points de friction quantitativement, cette méthode permet également d'identifier rapidement les problèmes dans le parcours du consommateur et d'offrir des solutions pour optimiser son parcours en ligne. Finalement, cette représentation visuelle permet de comparer différents parcours, afin d'identifier les ressemblances et différences, tant au niveau de la durée des différentes sous-tâches que de l'emplacement des différents points de friction. Par exemple, comparer le nombre de points de friction de différents compétiteurs dans la sous-tâche paiement permet d'avoir une meilleure idée de la performance relative de chacun pour cette sous-tâche.

De plus, les résultats présentés dans le chapitre 2 démontrent que les utilisateurs ne sont pas en mesure d'identifier rétrospectivement avec précision les points de friction dans leur propre parcours. En effet, les résultats obtenus dans cette étude démontrent que moins de 25% des points de friction implicites ont été identifiés par les participants lors de leur évaluation rétrospective. Par ailleurs, de façon surprenante, 9% des points de friction ont été identifiés comme des éléments positifs de leur expérience par les participants, alors que les résultats des mesures psychophysiologiques démontrent le contraire. Ces résultats sont assez consistants entre les différents sites d'épicerie en ligne testés, ce qui augmente la généralisation des résultats.

Dans un deuxième temps, ce mémoire avait également pour but de répondre aux questions suivantes :

Q3 : Quel est le rôle de la satisfaction dans la relation entre l'expérience utilisateur du consommateur et ses intentions ?

Q4 : Quel est l'impact d'une expérience utilisateur négative sur la satisfaction du consommateur ?

Celles-ci ont été répondues dans l'article 3 en intégrant les résultats de la méthode présentée précédemment. Puisque le but était d'établir des liens entre les mesures psychophysiologiques implicites et les mesures implicites de plusieurs construits, dont la satisfaction, une seconde collecte identique a été effectuée afin d'augmenter la taille de l'échantillon. Sans présenter visuellement les résultats, nous avons utilisé la méthode de calcul des points de friction psychophysiologiques implicites afin de calculer le nombre de points de friction dans la sous-tâche « magasinage » pour chacun des participants. Ce construit, appelé « émotions négatives » a été utilisé afin d'évaluer l'impact d'une expérience utilisateur négative sur la satisfaction du consommateur. Afin de répondre plus précisément aux questions Q3 et Q4, quatre hypothèses ont été formulées, basées sur la littérature existante.

L’Hypothèse 1 (H1) postulait que les intentions de réachat allaient être influencées positivement par la satisfaction perçue en ligne et hors ligne. Plus spécifiquement, les intentions de réachat allaient être influencées par la qualité du design de l’interface web (H1a), la sécurité perçue en ligne (H1b), la réalisation de la commande hors ligne (H1c) ainsi que le service à la clientèle hors ligne (H1d). Les résultats de l’étude ont supporté positivement les quatre hypothèses.

L’Hypothèse 2 (H2) postulait que les intentions de recommander le site internet allaient être influencées positivement par la satisfaction perçue en ligne et hors ligne. Plus spécifiquement, les intentions de recommander le site internet allaient être influencées par la qualité du design de l’interface web (H2a), la sécurité perçue en ligne (H2b), la réalisation de la commande hors ligne (H2c) ainsi que le service à la clientèle hors ligne (H2d). Les résultats de l’étude ont supporté positivement les quatre hypothèses.

L’Hypothèse 3 (H3) énonçait que les émotions négatives vécues dans la sous-tâche « magasinage » allaient influencer la satisfaction perçue en ligne ainsi que les intentions, Plus spécifiquement, avoir une expérience vécue négative allait affecter la satisfaction des consommateurs envers le design de l’interface web (H3a) et la sécurité perçue en ligne (H4b), ainsi que les intentions de réachat (H3c) et les intentions de recommencer le site internet (H3d). Les résultats de l’étude ont supporté les quatre hypothèses.

L’Hypothèse 4 (H4) postulait que la satisfaction perçue en ligne allait médier l’effet entre les émotions négatives vécues des consommateurs et leurs intentions. Plus spécifiquement, la satisfaction perçue envers le design de l’interface web allait médier la relation entre les émotions négatives vécues et les intentions de réachat (H4a) et les intentions de recommander le site web (H4b). De façon similaire, la satisfaction perçue envers la sécurité en ligne allait médier la relation entre les émotions négatives vécues et les intentions de réachat (H4c) et les intentions de recommander le site web (H4d). Les résultats de l’étude ont supporté deux des quatre hypothèses, démontrant que la satisfaction perçue envers le design de l’interface web médie la relation entre les émotions négatives vécues et les intentions de réachat (H4a) et les intentions de recommander le site web (H4b). Cependant, la satisfaction perçue envers la sécurité en ligne ne semble pas jouer un rôle médiateur dans la relation entre les émotions négatives vécues et les intentions de réachat (H4b) et les intentions de recommander le site web (H4d).

Ces résultats ont donc permis de montrer que la satisfaction perçue en ligne, plus précisément la satisfaction perçue envers le design de l'interface web, joue un rôle crucial dans la relation entre l'expérience du consommateur et ses intentions, en venant médier la relation. De plus, les résultats de l'étude ont permis de conclure que l'expérience client négative vécue a un impact significatif sur la satisfaction en ligne ainsi que les intentions.

Contributions

D'un point de vue théorique, les résultats contribuent à la littérature existante dans plusieurs domaines. Alors que les chapitres 2 et 3 contribuent à la littérature en expérience utilisateur, le chapitre 4, lui, apporte des contributions significatives aux domaines du marketing et du commerce électronique.

D'un côté, les résultats de l'étude contribuent au domaine de l'expérience utilisateur en proposant une méthode de visualisation fiable permettant de visualiser les points de friction psychophysiologiques implicites dans le parcours du consommateur. En introduisant la notion de points de friction psychophysiologiques implicites, cette étude permet d'identifier des points de friction auparavant non détectés par les méthodes précédemment utilisées, comme les questionnaires et les entrevues, (Fang et al., 2014) et de fournir des informations moins biaisées et donc, plus crédibles (de Guinea et al., 2014). Ainsi, cette méthode permet d'obtenir plus de précision et de fiabilité dans l'identification des points de friction, comparé aux points de friction mentionnés rétrospectivement par les consommateurs après leur interaction.

D'un autre côté, les résultats de l'étude combinant les mesures psychophysiologiques implicites et explicites contribuent aux domaines du marketing et du commerce électronique de plusieurs façons.

Dans un premier temps, en étudiant le processus d'achat omnicanal complet, des attentes préachat à la satisfaction post collecte de la commande en épicerie, les résultats permettent d'obtenir une meilleure compréhension de la formation de la satisfaction en ligne et hors ligne, ce qui n'a jamais été fait dans ce contexte dans la littérature antérieure (Kim, Ferrin & Rao, 2009).

Dans un deuxième temps, les résultats obtenus suggèrent qu'une expérience négative vécue, mesurée à l'aide de la méthode des points de friction psychophysiologiques implicites, affecte non seulement la satisfaction perçue en ligne, mais également les intentions des consommateurs. Ces résultats, cohérents avec la littérature existante sur la satisfaction (Dube-Rioux, 1990, Price et al., 1995; Oliver, 1993, Hui and Tse, 1996) et les intentions (Inman et al., 1997; Zeelenberg & Pieters, 2004), démontrent que les émotions négatives perçues influencent la satisfaction et les intentions, mesurées séparément. Cependant, les résultats contribuent à la littérature en montrant la différence entre les émotions négatives vécues et perçues. Ils montrent aussi l'impact que peut avoir l'expérience vécue en ligne sur les intentions, qui sont mesurées, souvent jusqu'à une semaine plus tard, suite à la collecte de la commande en épicerie. Ainsi, en intégrant des construits implicites à des construits empiriquement testés à maintes reprises dans la littérature en marketing, comme la satisfaction et les intentions (Oliver, 1997; Oliver, 1993), cette étude contribue à obtenir une meilleure compréhension de la façon dont les construits implicites et explicites interagissent entre eux dans la formation de la satisfaction et des intentions (De Guinea et al., 2014).

Dans un troisième temps, en identifiant le rôle médiateur de la satisfaction perçue envers le design de l'interface web dans la relation entre les émotions négatives vécues et les intentions de réachat et les intentions de recommander le site web, les résultats démontrent que la satisfaction perçue face à l'expérience n'est pas à négliger au profit de l'expérience implicite vécue. Par conséquent, cette étude contribue à démontrer la complémentarité des mesures explicites et implicites. En effet, les émotions perçues des consommateurs sont toutes aussi importantes que celles réellement vécues en ce qui concerne les intentions des consommateurs. Afin d'avoir un portrait complet et non biaisé de l'expérience, il est important de toutes les considérer.

Implications

D'un point de vue pratique, cette étude propose un nouvel outil pour le domaine de l'expérience utilisateur. Celui-ci permettra aux professionnels de l'industrie d'identifier plus facilement et plus rapidement des points de friction implicites dans le parcours de l'utilisateur, des facteurs très importants à considérer dans le choix d'outils d'évaluation en UX (Georges et al., 2017). Par ailleurs, cette nouvelle méthode, à l'aide de la combinaison des mesures psychophysiologiques, permet d'obtenir des résultats beaucoup plus précis, permettant aux professionnels souhaitant améliorer une interface d'avoir une liste des moments précis où l'utilisateur a vécu une émotion négative intense. Finalement, cette méthode peut également être très utile pour comparer entre eux des tâches, des compétiteurs ou encore, des versions futures et antérieures. Cette étude, de par sa grande validité écologique, permet également aux épiciers d'obtenir une meilleure compréhension des attentes et de la satisfaction des consommateurs dans un contexte réel. En effet, rares sont les études qui étudient le processus du début à la fin et requièrent que le participant magasine avec sa propre carte de crédit, selon ses besoins réels. Il a été prouvé que les intentions des consommateurs diffèrent souvent de leurs actions réelles (Venkatesh et al., 2008; Sheeran et al., 1999; Landis et al., 1978). Par conséquent, le design expérimental de cette étude permet d'obtenir une plus grande richesse d'information. Dans le même ordre d'idées, rares sont les études utilisant plusieurs canaux, forçant ainsi les consommateurs à interagir à la fois avec l'épicier en ligne et hors ligne dans le même parcours, rendant la comparaison entre les deux à la fois facile et nécessaire, puisque les consommateurs s'attendent au même niveau de service au travers des deux canaux.

Limites et recherches futures

Afin de mettre en contexte les résultats obtenus, il est nécessaire de reconnaître certaines limites de l'étude. Tout d'abord, la méthode de visualisation développée et testée dans ce mémoire a pour l'instant seulement été appliquée au contexte de l'épicerie en ligne, de nature utilitaire. Or, un contexte utilitaire génère habituellement peu de variation au niveau de l'intensité émotionnelle ressentie par le consommateur, ce qui n'est plus le cas dans un contexte plus hédonique, comme celui des jeux vidéo. Par conséquent, tester cette méthode dans un contexte hédonique contribuerait à la généralisation des résultats.

Dans un deuxième temps, l'expérience était assez longue comparée à une expérience en ligne moyenne. Par conséquent, il est possible que certains points de friction implicites soient liés à la fatigue des participants, obligés de terminer la tâche, plutôt qu'à un problème généré par le site internet. Pour réduire cette limite, d'autres études de plus courtes durées pourraient être réalisées.

Dans un troisième temps, cette étude a été réalisée avec 21, puis 45 participants, ce qui peut être considéré comme un échantillon de petite ou moyenne taille dans certains domaines. Cette taille d'échantillon est cependant assez courante en expérience utilisateur, où les études utilisant des données psychophysiologiques ont en moyenne entre 40 et 50 participants (Riedl et al., 2009). Cette étude pourrait être répliquée avec un plus grand nombre de participants afin d'augmenter la validité des résultats obtenus. De plus, l'échantillonnage effectué est un échantillonnage principalement de convenance, puisque les participants ont été recrutés via le Panel HEC Montréal. Bien que les jeunes adultes représentent une grande partie de la population en ligne (Statista, 2018), il serait intéressant de reproduire l'expérience avec des groupes d'âge plus diversifiés, par exemple en intégrant des consommateurs de plus de 50 ans et en comparant les résultats obtenus.

Finalement, dans le cadre de l'étude, les participants ont été assignés aléatoirement à un site d'épicerie en ligne. Cependant, cette assignation diminue la validité écologique de l'étude et peut affecter la satisfaction des participants, puisque l'épicier assigné n'est pas nécessairement le préféré des participants, pour plusieurs raisons comme la proximité entre l'épicerie et le domicile du participant. Par conséquent, des projets futurs pourraient contrôler pour ces facteurs en effectuant une présélection des participants.

Bibliographie

Alexandris, K, S Douka, P Papadopoulos et A Kaltsatou (2008). « Testing the role of service quality on the development of brand associations and brand loyalty », *Managing Service Quality: An International Journal*, vol. 18, no 3, p. 239-254.

Anesbury, Zachary, Magda Nenycz-Thiel, John Dawes et Rachel Kennedy (2016). « How do shoppers behave online? An observational study of online grocery shopping », *Journal of Consumer Behaviour*, vol. 15, no 3, p. 261-270.

Ariely, Dan (1998). « Combining experiences over time: the effects of duration, intensity changes and on-line measurements on retrospective pain evaluations », *Journal of Behavioral Decision Making*, vol. 11, no 1, p. 19-45.

Baumeister, Roy F, Ellen Bratslavsky, Catrin Finkenauer et Kathleen D Vohs (2001). « Bad is stronger than good », *Review of general psychology*, vol. 5, no 4, p. 323-370.

Bloomberg (2019). « Online food and beverage market share in the United States in 2016, by company. dans *Statista - The Statistics Portal*. Récupéré le 28 février 2018

Citrin, Alka Varma, Donald E Stem Jr, Eric R Spangenberg et Michael J Clark (2003). « Consumer need for tactile input: An internet retailing challenge », Journal of Business research, vol. 56, no 11, p. 915-922.

Cockburn, Andy, Philip Quinn et Carl Gutwin (2017). « The effects of interaction sequencing on user experience and preference », *International Journal of Human-Computer Studies*, vol. 108, p. 89-104.

De Guinea, Ana Ortiz, Ryad Titah et Pierre-Majorique Léger (2013). « Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research », *Computers in Human Behavior*, vol. 29, no 3, p. 833-844.

De Guinea, Ana Ortiz, Ryad Titah et Pierre-Majorique Léger (2014). « Explicit and implicit antecedents of users' behavioral beliefs in information systems: A neuropsychological investigation », *Journal of Management Information Systems*, vol. 30, no 4, p. 179-210.

Desrochers, Camille.; Léger, Pierre-Majorique.; Mirhoseini, SeyedMohammadmahdi.; Fredette, Marc.; Sénécal, Sylvain (Forthcoming). « The Arithmetic Complexity of Online Grocery Shopping: The Moderating Role of Product Pictures. » *Industrial Management & Data Systems.*

Dube-Rioux, Laurette (1990). « The power of affective reports in predicting satisfaction judgments », ACR North American Advances.

Fang, Yulin, Israr Qureshi, Heshan Sun, Patrick McCole, Elaine Ramsey et Kai H Lim (2014). « Trust, satisfaction, and online repurchase intention: The moderating role of perceived effectiveness of e-commerce institutional mechanisms », *Mis Quarterly*, vol. 38, no 2.

Forrester (2018). « Online Grocery Retail Forecast, 2017 to 2022 (Global) », dans *Forrester Report*, Récupéré le 15 juin 2018

Huang, Yan & Oppewal, Harmen. (2006). « Why consumers hesitate to shop online: An experimental choice analysis of grocery shopping and the role of delivery fees. *International Journal of Retail & Distribution Management*, vol 34, no 4/5, p. 334-353.

Hui, Michael K et David K Tse (1996). « What to tell consumers in waits of different lengths: An integrative model of service evaluation », *Journal of Marketing*, vol. 60, no 2, p. 81-90.

Inman, J Jeffrey, James S Dyer et Jianmin Jia (1997). « A generalized utility model of disappointment and regret effects on post-choice valuation », *Marketing Science*, vol. 16, no 2, p. 97-111.

Kahn, Barbara E, J Jeffrey Inman et Peter C Verhoef (2018). « Introduction to Special Issue: Consumer Response to the Evolving Retailing Landscape », *Journal of the Association for Consumer Research*, vol. 3, no 3, p. 255-259.

Kahneman, Daniel, Barbara L Fredrickson, Charles A Schreiber et Donald A Redelmeier (1993). « When more pain is preferred to less: Adding a better end », *Psychological science*, vol. 4, no 6, p. 401-405.

Kim, Dan J (2012). « An investigation of the effect of online consumer trust on expectation, satisfaction, and post-expectation », *Information Systems and E-Business Management*, vol. 10, no 2, p. 219-240.

Kim, Dan J., Donald L. Ferrin et H. Raghav Rao (2009). « Trust and Satisfaction, Two Stepping Stones for Successful E-Commerce Relationships: A Longitudinal Exploration », *Information Systems Research*, vol. 20, no 2, p. 237-257.

Landis, Dan, Harry C Triandis et John Adamopoulos (1978). « Habit and behavioral intentions as predictors of social behavior », *The Journal of Social Psychology*, vol. 106, no 2, p. 227-237.

Lee, Hyung Seok (2013). « Major moderators influencing the relationships of service quality, customer satisfaction and customer loyalty », *Asian Social Science*, vol. 9, no 2, p. 1.

Lourties, Sébastien, Léger, Pierre-Majorique, Sénécal, Sylvain, Fredette, Marc et Shang Lin Chen. (2018) « Testing the Convergent Validity of Continuous Self-Perceived Measurement Systems: An Exploratory Study ». In : *International Conference on HCI in Business, Government, and Organizations*. Springer, Cham, 2018. p. 132-144.

McNair, Corey (2018). « Worldwide retail and ecommerce sales: Emarketer's updated forecast and new mcommerce estimates for 2016—2021»,dans *eMarketer*, 11 p. Récupéré le 2 avril 2019

Myers, James H (1991). « Measuring customer satisfaction: is meeting expectations enough? », *Marketing Research*, vol. 3, no 4.

Nielson (2017). « What's in-store for Online Grocery Shopping, Omnichannel Strategies to Reach Crossover Shoppers. » dans *Nielson Global Connected Commerce Report*, Récupéré le 15 juin 2018

Oliver, R. L. "A Cognitive Model for the Antecedents and Consequences of Satisfaction," Journal of Marketing Research (17), 1980, pp. 460-469
Oliver, R. L. (1997). Satisfaction: A behavioral perspective on the consumer.

Oliver, Richard L. (1993). « Cognitive, Affective, and Attributes Bases of the Satisfaction Response », *Journal of Consumer Research*, vol. 20, no 3, p. 418-430.

Price, Linda L, Eric J Arnould et Patrick Tierney (1995). « Going to extremes: Managing service encounters and assessing provider performance », *Journal of marketing*, vol. 59, no 2, p. 83-97.

Riedl, René, Thomas Fischer et Pierre-Majorique Léger (2017). « A decade of NeuroIS research: status quo, challenges, and future directions ».

Salesforce Research. (n.d.). « Retail website visit duration in the United States as of 4th quarter 2018, by device (in minutes) » dans *Statista - The Statistics Portal*. Récupéré le 2 avril 2019

Schooler, Jonathan W et Erich Eich (2000). « Memory for emotional events ».

Sheeran, Paschal, Paul Norman et Sheina Orbell (1999). « Evidence that intentions based on attitudes better predict behaviour than intentions based on subjective norms », *European journal of social psychology*, vol. 29, no 2-3, p. 403-406.

Statista (2017). « U.S. consumers: Online Grocery Shopping-Statistics & Facts », dans *Statista - The Statistics Portal*., Récupéré le 20 juin 2018

The Hartman Group & FMI (2018). « Share of shoppers who have used an online-only retailer for groceries in the United States in 2018, by generation » . dans *Statista - The Statistics Portal*. Récupéré le 28 février 2019

Venkatesh, Viswanath, Susan A Brown, Likoebe M Maruping et Hillol Bala (2008). « Predicting different conceptualizations of system use: the competing roles of behavioral intention, facilitating conditions, and behavioral expectation », *MIS quarterly*, p. 483-502.

Zeelenberg, Marcel et Rik Pieters (2004). « Beyond valence in customer dissatisfaction: A review and new findings on behavioral responses to regret and disappointment in failed services », *Journal of business Research*, vol. 57, no 4, p. 445-455.

Annexes

Annexe 1: Interview Guide

Part 1: Perception of the process

1. Overall, how would you compare online groceries to traditional (in-store) groceries?

2. What would be the main positives points in your experience of the process?

3. What would be the main points to improve in your experience of the process?

4. How did you find the process of choosing where to collect your order?

5. How did you find the process of choosing when to collect your order?

6. How did you find the shopping process (search and product selection)?

7. How did you find the account creation process?

8. How did you find the payment process?

Part 2: Time aspect

9. What did you think about the time needed to complete all the steps?

Part 3: General Impressions

10. What is your impression of the security of the information required to complete the process?

11. What do you think about the distance required to pick up your collection?

12. Do you think you will redo online grocery shopping in your daily life? Why?

13. Would you recommend this online grocery store?