[Page de garde]

HEC MONTRÉAL

Impact de la segmentation de tutoriels vidéos sur l'expérience vécue et perçue des utilisateurs en contexte de formation en technologies de l'information

par

Charles Lamontagne

Sciences de la gestion
(Option Expérience utilisateur en contexte d’affaires)

Mémoire présenté en vue de l’obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)

Décembre 2019

HEC MONTRÉAL

Retrait d'une ou des pages pouvant contenir des renseignements personnels

Résumé

La segmentation des tutoriels vidéos, soit le découpage plus fin des instructions, en contexte de formation en technologie de l'information a été démontrée comme bénéfique pour les apprenants sur plusieurs aspects. Aucune recherche ne s'est toutefois penchée sur l'impact de la segmentation des tutoriels vidéos sur l'expérience vécue et perçue des utilisateurs lorsqu'ils interagissent avec la technologie cible de la formation. Les émotions et les perceptions ont un impact important sur l'évaluation que portent les utilisateurs au premier contact d'une nouvelle technologie, qui se fait généralement lors de la formation utilisateur pour les employés. Ceci démontre le besoin de comprendre l'expérience qu'ils vivent au cours de la formation utilisateur.

Ce mémoire par articles étudie l'impact de la segmentation des tutoriels vidéos sur l'expérience vécue et perçue. Une étude en laboratoire a été conduite avec 2 groupes de 10 participants afin d'étudier cette question, un groupe avec des tutoriels segmentés et l'autre avec des tutoriels non segmentés. L'expérience vécue a été évaluée à l'aide mesures implicites (psychophysiologiques) et l'expérience perçue a été évaluée à l'aide de mesures explicites (questionnaires). Les résultats obtenus montrent une différence dans l'expérience perçue des participants, mais pas dans l'expérience vécue. Le projet de recherche a également mené à une première ligne directrice pour le nombre de participants nécessaire pour la réalisation de test utilisateur avec des outils de mesures psychophysiologiques. Un chapitre de ce mémoire y est dédié.

Cette étude contribue à la recherche en comblant le manque de recherche sur l'impact de la segmentation des tutoriels vidéos sur l'expérience vécue et perçue des apprenants en contexte de formation utilisateur. Ce mémoire a contribué à développer une nouvelle méthodologie pour tester les vidéos de formation permettant d'obtenir des rétroactions riches et rapides.

Mots clés : expérience utilisateur, point de friction psychophysiologique, test utilisateur, mesures psychophysiologiques, formation utilisateur, tutoriel vidéo

Table des matières

Liste des tableaux et des figures

Liste des tableaux

Liste des figures

Table 1. Contributions et responsabilités personnelles 7

Table 2. Details of psychophysiological pain points experienced per participant 19

Table 3. Operationalization of Constructs 45

Table 4. Task completion time, PPPs and SAM scale mean results for Long task and Short task 48

Table 6. NASA-TLX mean scores for the 2 conditions 49

Table 7. SUS mean score for the 2 conditions 49

Liste des abréviations

EDA : «*Electrodermal activity* »

NASA-TLX : *«NASA Task Load Index*»

PPP : Point de friction psychophysiologique / «*Psychophysiological pain point* »

ROI : «*Return on investment* »

SAM scale : *«Self-Assessed Manikin scale*»

SI / IS : Système d'information / «*Information System*

SUS : «*System Usability scale*»

TAM : «*Technology Acceptance Model* »

TI / IT : Technologie de l'information / «*Information Technology* »

UP : «*Usability problem* »

UTAUT : «*Unified Theory of Acceptance and Use of Technology* »

UX : Expérience utilisateur / «*User Experience* »

Avant-propos

Suite à l'autorisation de la direction du programme de M.Sc. en Gestion de HEC Montréal, ce mémoire a été rédigé sous la forme de deux articles. Tous les coauteurs ont donné leur consentement pour que ces articles soient utilisés pour ce mémoire. Le comité d'éthique en recherche (CER) de HEC Montréal a approuvé la collecte de données pour ce projet de recherche.

Remerciements

« *Tous les chemins mènent à Rome* »

Je tiens à remercier mes directeurs de mémoire, Pierre-Majorique Léger et Sylvain Sénécal, de m'avoir confié ce projet et de m'avoir permis de le mener à bien de façon autonome.

Je remercie également l'équipe de Deloitte Digital, dont Romain Pourchon, pour ce projet stimulant qui m'a préparé à toute éventualité et poussé à résoudre des contingences relatives aux implantations de systèmes d'information organisationnels.

J'aimerais remercier tout spécialement tous les membres de l'équipe du Tech3Lab qui m'ont épaulé de près ou de loin au cours des différentes phases de mon projet : Emma, Bertrand, Robin, David, Frédérique, Shang Lin ainsi que tous les assistant(e)s de recherche.

Merci à mes amis, camarades et collègues qui ont facilité mon établissement dans la grande métropole : Ariane R., Claudel, Ariane B., Caroline, Marianne, Audrey, Catherine, Félix, Éric, Carol-Ann, Mario et tous les autres.

Josée, Bruno, qui l'eut cru, ça y est enfin, la fin de ce périple qu'a été mon parcours universitaire. Merci pour votre appui au cours de ces sept années et demie.

Finalement, merci Alexis de m'avoir encouragé et soutenu même dans les moments plus difficiles.

Introduction

Mise en contexte

La formation des utilisateurs est une étape cruciale pour la réussite des projets d'implantation de systèmes d'information (Gupta & Anson, 2014; Igbaria et al., 1995). Puisqu'il s'agit généralement du premier contact des employés avec le nouveau système d'information, il s'agit d'une étape cruciale de la gestion du changement qui a un impact significatif sur l'adoption des technologies (Igbaria et al., 1995; Gupta et al., 2010). Il est donc très important que la formation des utilisateurs soit optimisée pour les utilisateurs; ceux-ci doivent bien comprendre les procédures d'utilisation du nouveau système d'information et les avantages qui en découlent (Van der Meij & Van Der Meij, 2013, Gupta et al., 2010; Bostrom et al., 1990).

Les tutoriels vidéos sont désormais utilisés dans de nombreuses organisations pour effectuer de la formation utilisateur de systèmes d'information (Giannakos, 2013). Il s'agit de moyen efficace pour effectuer la formation des utilisateurs, et ce, de façon standardisée (Van Der Meij & Van Der Meij, 2013). L'un des principaux avantages de l'utilisation des tutoriels vidéos réside dans le fait que les utilisateurs peuvent regarder ces tutoriels de façon autonome, sans formateur, au moment qui est le plus opportun pour eux (Van Der Meij & Van Der Meij, 2013). Ceci implique toutefois que les tutoriels vidéos doivent être optimisés pour que les utilisateurs comprennent leurs contenus sans avoir accès à un formateur qui peut offrir des clarifications (Tversky et al., 2002; Shea & Bidjerano, 2010). D'ailleurs, les formateurs d'expérience sont généralement en mesure de détecter l'état psychologique de leurs apprenants et d'adapter leur formation en fonction, par exemple, de l'état de fatigue ou de fermeture des élèves (D'Mello et al., 2006; Graesser et al., 2005; Lepper & Woolverton, 2002). Les tutoriels vidéos sont standardisés et présentent les instructions de la même façon, quel que soit l'état de l'apprenant; ils doivent ainsi être en mesure de motiver les utilisateurs en plus de leur permettre de comprendre l'utilisation du nouveau système d'information (Tversky et al., 2002; Shea & Bidjerano, 2010).

Le design des tutoriels vidéos requiert une certaine expertise puisque plusieurs facteurs doivent être pris en compte pour que les utilisateurs en retirent des apprentissages (Kennedy, 2004; Mestre, 2012; Wouters et al., 2007). Entre autres, le débit des instructions est très important, un rythme trop lent peut être trop ennuyeux pour les utilisateurs et un rythme trop rapide peut saturer les utilisateurs, causant une dégradation des apprentissages (Bovair & Kieras, 1991; Linek et al., 2010).

Des tutoriels qui sont trop longs, qui possèdent une densité informationnelle trop élevée ou qui présentent des instructions complexes peuvent causer des problèmes d'apprentissage (Mayer & Moreno, 2003). La segmentation des tutoriels vidéos est une méthode employée pour remédier à ces situations (Spanjers et al., 2010). La segmentation des tutoriels vidéos est définie comme étant un découpage plus fin des instructions présentées comparé à un bloc d'informations présentées en continu (Spanjers et al., 2010). Plusieurs études se sont penchées sur les avantages de la segmentation des tutoriels vidéos (voir Mayer, 2005; Spanjers et al., 2010; Spanjers et al., 2012; Khacharem et al., 2013; Zacks & Tversky, 2003; Mayer & Moreno, 2003; Moreno, 2007; Zack et al., 2007). Toutefois, aucune recherche ne s'est intéressée à l'impact de la segmentation des tutoriels vidéos sur l'expérience vécue et perçue des utilisateurs lorsqu'ils interagissent avec le système d'information cible après avoir visionné ces tutoriels.

Plusieurs recherches suggèrent que les émotions et les perceptions des utilisateurs influencent l'évaluation que les utilisateurs portent sur une nouvelle technologie (Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015). Plusieurs facteurs cognitifs et affectifs impacteraient spécifiquement la facilité d'utilisation perçue et l'utilité perçue (Agarwal & Karahanna, 2000; Barki et al., 2008; Karahanna & Straub, 1999). L'expérience perçue est généralement évaluée à l'aide de mesures autodéclarées, qui sont des mesures explicites, à la fin de l'interaction

(Bruun et al., 2016). Ces méthodes comprennent les entrevues, les questionnaires et les données d'observation (Moon et al., 2016; Nenonen et al., 2018; Rawson et al., 2013). Il existe de nombreux questionnaires pour mesurer différents aspects des interactions humain-machine, tels que l'utilisabilité (Brooke, 1996; Loiacono et al., 2002), la charge de travail (Hart & Staveland, 1988) ainsi que la valence émotionnelle, activation émotionnelle et la dominance (Bradley & Lang, 1994).

Toutefois, de précédentes recherches ont démontré qu'il y a une différence entre ce que vivent les utilisateurs et ce qu'ils se souviennent d'avoir vécu au cours de leur interaction avec un système (Cockburn et al., 2017; Bruun et al., 2016; Hetland et al., 2018). Les évaluations rétrospectives seraient biaisées puisque la mémoire humaine est influencée par des moments plus marquants que d'autres au cours de l'expérience, qu'ils soient positifs ou négatifs (Cockburn et al., 2017; Bruun et al., 2016; Hetland et al., 2018). Les mesures explicites et observationnelles ne capturent donc pas toute la richesse des émotions ressenties par les utilisateurs au cours de leurs interactions (de Guinea et al., 2014). Il y a donc une perte d'information précieuse, puisque l'expérience vécue par l'utilisateur à chaque moment est dépendante de son état physiologique et psychologique à ce moment précis (Bruun et al., 2016; Mashapa et al., 2013). La littérature suggère que la combinaison de méthodes perceptuelles (explicites) et psychophysiologiques (implicites) permet d'avoir une compréhension plus profonde de l'expérience des utilisateurs (de Guinea et al., 2014; de Guinea et al., 2012).

Les mesures psychophysiologiques permettent d'évaluer l'état affectif et émotif des utilisateurs au cours de leur interaction avec une interface de façon implicite qui n'obstrue pas leurs mouvements (de Guinea et al., 2014). Ces mesures collectent en temps réel les manifestations physiologiques des utilisateurs qui proviennent de leur état psychologique (Andreassi, 2000). La superposition de plusieurs mesures psychophysiologiques permet d'avoir une compréhension en profondeur de l'état affectif et émotionnel des utilisateurs (de Guinea et al., 2014; Ganglbauer et al., 2009; Maia & Furtado, 2016).

La valence émotionnelle et l'activation émotionnelle sont deux mesures psychophysiologiques qui peuvent être recueillies lors de tests utilisateurs (Russel, 1980; Dawson, 2007; Ekman & Friensen, 1978; Ganglbauer et al., 2009; Maia & Furtado, 2016). La valence émotionnelle est une valeur se situant sur un intervalle de 1 à -1 représentant le caractère émotionnel plaisant (valence positive) ou déplaisant (valence négative) que vit l'utilisateur (Russel, 1980). L'activation émotionnelle permet de voir le degré de stimulation de l'utilisateur sur un intervalle entre très calme à très excité (Russel, 1980).

Les points de friction émotionnels et cognitifs permettent d'avoir une représentation approfondie des émotions négatives vécues par les utilisateurs au cours du parcours utilisateur (Giroux-Huppé et al., 2019). Les points de friction émotionnels sont définis par une haute valence négative et une haute activation émotionnelle de l'utilisateur qui sont causées par des moments non optimaux dans leurs interactions avec une interface (Giroux-Huppé et al., 2019). Les points de friction cognitifs sont définis par une haute charge cognitive et une valence négative. L'utilisation des mesures physiologiques permet d'identifier avec précision les moments où ces réactions se produisent dans l'interaction, ce qui donne des résultats plus fiables que ce que les utilisateurs sont capables d'identifier (Giroux-Huppé et al., 2019).

Objectifs de l'étude et questions de recherche

Ce mémoire a comme objectif d'étudier l'effet de la segmentation des tutoriels vidéos en contexte de formation en technologies de l'information sur l'expérience vécue, à l'aide de mesures implicites, et l'expérience perçue des utilisateurs, à l'aide de mesures explicites. La méthode d'identification des points de friction psychophysiologiques dans le parcours utilisateur a été développée récemment (Giroux-Huppé et al., 2019). En conséquence, contrairement aux méthodes traditionnelles utilisées en test utilisateur (Lewis, 2006; Nielsen & Landauer 1993; Nielsen 1994; Schmettow, 2012), cette méthode n'a pas une littérature exhaustive concernant le nombre d'utilisateurs nécessaires pour avoir des résultats pertinents et significatifs.

Ainsi, dans un premier temps, l'objectif de ce mémoire est de connaître le nombre d'utilisateurs minimal afin d'identifier une majorité de points de friction psychophysiologiques en test utilisateur. Dans un deuxième temps, après avoir trouvé le nombre d'utilisateurs nécessaires, il sera possible de l'utiliser pour comparer l'expérience vécue et perçue des utilisateurs selon le visionnement d'une version segmentée ou non segmentée de tutoriels vidéos. Ainsi, les questions de recherche de ce mémoire par articles se divisent de la façon suivante :

- Article 1 : Quel est le nombre minimal d'utilisateurs nécessaires pour identifier au moins 80% des points de friction psychophysiologiques dans un parcours utilisateur ?

- Article 2 : Quel est l'impact de la segmentation des tutoriels vidéos sur l'expérience implicite et explicite des utilisateurs ?

Information sur les articles

Article 1

Le premier article a été accepté et présenté à la *International Conference on Human Interaction & Emerging Technologies 2019* à Nice.

Résumé de l'article 1

La méthode d'identification des points de friction psychophysiologiques utilise des mesures physiologiques des utilisateurs plutôt que des mesures explicites et observationnelles qui sont généralement utilisés en tests utilisateurs. Cette méthode ayant été développée récemment, il n'y a pas de ligne directrice concernant le nombre minimal de participants pour la conduite de tests utilisateurs avec cette méthode. Ainsi, l'objectif de cette recherche est d'explorer combien d'utilisateurs sont nécessaires pour identifier au moins 80% des points de friction psychophysiologiques vécus au cours d'interactions humain-machine. Une expérience a été conduite en contexte de formation utilisateur. Les participants devaient regarder 5 vidéos de formation et exécuter les tâches qui leur étaient présentées après chaque vidéo. La valence émotionnelle, l'activation émotionnelle et la charge cognitive des participants (n = 15) ont été respectivement mesurées par la détection des émotions faciales, l'activité électrodermale et la pupillométrie. La valence émotionnelle, l'activation émotionnelle et la charge cognitive ont été triangulées pour obtenir les points de friction psychophysiologiques des participants. Ces moments où les utilisateurs ont vécu des émotions négatives ont été jumelés avec leur emplacement dans le parcours idéal, soit celui qui était montré dans les vidéos, afin d'avoir une base commune entre les participants permettant de les comparer pour trouver la première occurrence de chaque point de friction et de calculer leur nombre d'occurrences respectif. Les résultats montrent que 82% des points de friction psychophysiologiques ont été identifiés après 9 des 15 participants. L'indépendance de l'ordre des participants a été prouvée en utilisant 1000 permutations de l'ordre des participants, aboutissant aux mêmes résultats. Cette étude fournit une première ligne directrice pour le nombre de participants nécessaire pour des tests utilisateurs avec des mesures psychophysiologiques.

Article 2

Le deuxième article est présentement en préparation pour soumission dans la revue *Computers in Human Behavior*. Ainsi, ce mémoire présente une version préliminaire de cet article.

Résumé de l'article 2

L'objectif de cette recherche est d'explorer l'effet de la segmentation des tutoriels vidéos sur l'expérience perçue et vécue des utilisateurs en contexte de formation utilisateur. Une étude reproduisant un contexte de formation utilisateur a été menée en laboratoire. Les participants ont été assignés à l'une des deux conditions de la même formation utilisateur : une version avec 5 tutoriels vidéos et tâches à exécuter (n = 10) ou une version avec 8 tutoriels et tâches (n = 10). Ces 8 tutoriels présentaient exactement les mêmes instructions, mais avaient une segmentation plus fine puisque 2 vidéos (et tâches) de la première condition ont été divisés en plusieurs parties. Les deux tâches segmentées étaient les sujets d'intérêt de cette étude, les trois autres tâches servaient à créer un environnement de formation utilisateur similaire pour les deux conditions. L'expérience vécue a été évaluée par des mesures psychophysiologiques au cours des interactions des participants avec le système d'information et l'expérience perçue a été évaluée par questionnaire après l'exécution des deux tâches segmentées ainsi qu'à la fin de l'expérimentation. Les mesures psychophysiologiques ont été utilisées pour identifier les points de friction psychophysiologiques des participants de chaque condition. Les résultats des questionnaires et le nombre de points de friction psychophysiologiques vécus ont été comparés entre les deux conditions. Les résultats suggèrent que malgré une expérience vécue similaire, la segmentation des tutoriels a un impact positif sur l'expérience perçue des participants par rapport à la non-segmentation des tutoriels. Plus spécifiquement, la segmentation a un effet sur l'utilisabilité perçue du logiciel ainsi que sur la charge mentale perçue et la frustration perçue. Cette étude contribue à la littérature de la formation utilisateur en illustrant les impacts de la segmentation des tutoriels vidéos sur l'expérience vécue et perçue des utilisateurs.

Contributions et responsabilités personnelles

Les articles composant ce mémoire ont été réalisés dans le cadre d'expériences au Tech3Lab de HEC Montréal. Le tableau ci-dessous décrit ma contribution dans les différentes étapes des processus de réalisation des expérimentations et de l'écriture des articles qui en découlent. Ma contribution y est inscrite en pourcentage.

Table 1. Contributions et responsabilités personnelles

Étape du processus	Contribution

Définition des besoins d'affaires du partenaire	Identification des besoins du partenaire industriel et transformation en question de recherche scientifique - 50% • Projet fourni par le partenaire industriel • L'équipe du Tech3Lab a contribué à la définition de la question de recherche à partir du projet existant et des besoins du partenaire industriel ainsi que l'approche à privilégier pour y répondre.
Revue de la littérature	Rédaction de la revue de littérature - 100 %
Conception du design expérimental	Rédaction de la demande au CER et des demandes de changements - 75% • Projet s'inscrivant dans la continuité d'un projet précédent du Tech3Lab. • L'équipe du Tech3Lab s'est assuré que les demandes complétées au CER soient remplies adéquatement. Conception des protocoles des expériences - 75 % • Le protocole de la première étude a été fait majoritairement par l'équipe du Tech3Lab. • Le protocole de la deuxième étude comprenant les 2 conditions a été adapté à partir de celui de la première étude.

	Création du matériel de formation pour les collectes - 90% • Les vidéos de formation qui ont été utilisés initialement pour les prétests étaient fournis par le partenaire industriel • J'ai produit les tutoriels pour la collecte de l'article 1 à partir des vidéos fournis précédemment. • J'ai produit les tutoriels pour la collecte de l'article 2 à partir de ceux de la collecte précédente en ajoutant certaines modifications.
Recrutement des participants	Recrutement des participants - 90% • Le recrutement de quelques participants s'est fait par la plateforme de recrutement.
Prétests et collectes de données	Responsable de la collecte lors des prétests pour les deux collectes - 100% **Article 1:** Responsable des opérations lors des collectes - 80% • L'équipe du Tech3Lab posait les instruments de mesure et donnait les instructions aux participants. • Supervision du bon déroulement des collectes et pour dépanner en cas de problème avec le système d'information.

	• La transformation des données du participant précédent était effectuée au cours du participant qui le suivait, ce qui n'était pas exécuté dans la même salle que la collecte. • Présence à presque toutes les collectes sauf quelques exceptions où un cours suivi coïncidait avec l'horaire de collecte. • Présence à distance pour assister lors de problèmes techniques en cas d'absence. **Article 2 :** Responsable des opérations lors des collectes - 80% • L'équipe du Tech3Lab posait les instruments de mesure et donnait les instructions aux participants. • Supervision du bon déroulement des collectes et pour dépanner en cas de problème avec le système d'information.
Extraction et transformation des données	Extraction des données de test - 100% Transformation des données - 100% Triangulation des résultats - 50% • L'équipe des opérations du Tech3Lab a aidé pour l'utilisation du logiciel permettant la triangulation des données.

Analyse des données	Analyse des points de friction psychophysiologiques - 100% • La description des points de friction a été réalisée à l'aide de l'équipe des opérations du Tech3Lab lors des prétests. Analyses statistiques - 65% • Aide par le statisticien du Tech3Lab pour les tests statistiques plus complexes. • Statistiques descriptives • Définition des besoins relatifs aux tests statistiques pour répondre aux questions de recherche • Vérification de la conformité des résultats Interprétation des résultats - 85% • Aide par le statisticien du Tech3Lab pour les tests statistiques plus complexes.
Rédaction des articles	Rédaction des deux articles scientifiques - 100% • Les articles ont été améliorés grâce à quelques commentaires des coauteurs tout au long de la rédaction.

Chapitre 1 : Premier article
User Test: How Many Users Are Needed to Find the Psychophysiological Pain Points in a Journey Map?

Charles Lamontagne[1], Sylvain Sénécal[1], Marc Fredette[1], Shang Lin Chen[1], Romain Pourchon[2], Yohan Gaumont[2], David De Grandpré[2], Pierre-Majorique Léger[1]

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

{charles.lamontagne, sylvain.senecal, marc.fredette, shang-lin.chen,

pierre-majorique.leger}@hec.ca

[2] Deloitte Digital, La Tour Deloitte, 1190, avenue des Canadiens-de-Montréal, Bureau 500, Montréal, Qc, Canada H3B 0M7

{rpourchon, ygaumont, dadegrandpre}@deloitte.ca

Abstract:

The objective of this research is to investigate how many users are needed in usability testing to identify psychophysiological pain points (PPP) experienced by users during a human-computer interaction (HCI). Fifteen subjects were tested in a new user training context and results show that out of 15 participants, 82% of the total PPP were experienced after 9 participants. Calculations using 1000 trials of random orders were then performed to demonstrate the independence of the order. This research provides guidelines about what could be an ideal sample size for user test using psychophysiological measures.

Keywords: User experience (UX) · Usability testing · Psychophysiological measures · Psychophysiological pain points · Pain points

1.1 Introduction

Within the field of user experience (UX), pain points are known as user irritants impacting their interaction with digital entities (Platzer, 2018). Identifying pain points can be crucial to improve the usability, satisfactory and adaptability of interactions (Dirigan & Göktürk, 2011). The majority of methods currently used in UX evaluation are based on explicit (i.e., self-reported) and observational data (Bruun et al., 2016). However, previous research suggests that traditional approaches may not capture users' reactions to a new interface in an unbiased way (de Guinea et al., 2014). Hence, psychophysiological measures can be an alternative method to comprehend users' emotional states (Dirigan & Göktürk, 2011). This paper is building upon previous work of Giroux-Huppé et al. (2019), which proposes a new method to analyze online user journeys by identifying psychophysiological pain points; hereinafter referred to as PPPs (Giroux-Huppé et al., 2019). PPPs are defined as interaction moments for which users experience high emotional arousal and high negative emotional valence, as well as moments were users experience high cognitive load and high negative emotional valence (Giroux-Huppé et al., 2019). Past research in HCI has provided guidelines for self-reported UX test sample size, but there is no evidence if these apply to PPPs. Therefore, this research focuses on investigating the optimal number of users needed in order to identify most PPPs in a given user journey.

1.2 Background

1.2.2 Sample Size in Usability Testing

Usability is the user's ability to use a system, product, or service to perform a given task successfully (Tullis & Albert, 2013). On the other hand, UX encompasses all aspects of the user's interaction with the entity, which extend to the emotional and cognitive states, thoughts, and perceptions of the user (Tullis & Albert, 2013). Thus, the main purpose of user testing is to evaluate the user interactions with a system, product, or service and to identify usability problems (UPs) and pain points within their individual journey, thereby providing feedback to designers in order to improve the system (Bruun et al., 2016).

As user tests can be very expensive, the benefit of each additional participant must surpass their marginal cost, which emphasizes the need for practitioners to have the optimal sample size (Lewis, 2006). Early research with the thinking aloud method suggests that 5 participants are usually necessary in order to obtain up to 77% to 85% of UPs (Nielsen & Landauer, 1993). Testing additional users increases the number of UPs found, but progressively decreases the marginal gain (Nielsen & Landauer, 1993). Since its publication, "The Magic Number 5" rule and sample size in user testing have induced a long-running debate that has not reached a consensus yet (Nielsen & Landauer, 1993; Nielsen, 1994; Schmettow, 2012). Nevertheless, this number is considered as the reference in usability testing (Tullis & Albert, 2013).

Moreover, lab-based user research methods, such as usability testing, usually aim to find and resolve problems (Rohrer, 2014). Hence, those methods analyze the users' behaviors with a more qualitative approach (Rohrer, 2014). Thereby, a smaller sample size suffices as it allows one to gain a deeper understanding of the participants' perception of the experience (Vernette et al., 2008).

1.2.3 Psychophysiological Measures in Usability Testing

Traditional methods to assess a user's experience throughout experimentation has been proven to render only a portion of the user's feelings toward a new digital product or service, as these methods rely only on explicit and observational measures (de Guinea et al., 2014). Furthermore, users may encounter UPs throughout a user test, but usually only convey their emotional reactions at the end of it, which does not allow one to fully understand each moment-to-moment user reaction (Bruun et al., 2016). Hence, research shows that self-reported measures represent more of an overall evaluation by the user, which has been proven to be influenced by the Peak-End Effects (Cockburn et al., 2015; Ganglbauer et al., 2009).

Psychophysiological measures provide an unobstructive and implicit manner to ascertain a user’s affective and cognitive conditions in HCI (Dirigan & Göktürk, 2011). This is a reliable approach to comprehend the user's emotional state in real time during usability testing, as it allows to record physiological manifestations of the user’s psychological sentiments (Andreassi, 2000). At least two measures need to be simultaneously used to have a rich comprehension of the user's state (Ganglbauer et al., 2009, Maia & Furtado, 2016). For example, facial emotions can be detected and analyzed to fit a range from negative to positive, i.e., emotional valence (Dawson, 2007). Arousal level can be measured with electrodermal activity (EDA), ranging from very calm to neutral to highly stimulated (Ekman & Friensen, 1978). Cognitive load can be measured with pupillometry, which is assessed with eye-tracking technologies, as they allow the measurement of the pupil dilatation (Paas et al., 2010; Sweller et al., 2011; Wang, 2011).

1.2.4 PPP

Pain point is a term rooted in the medical domain, in which it refers to the localization of a hostile sensation caused by unbearable, perturbing, or even harmful stimuli (Platzer, 2018). In UX, it is widely known as an emotional irritant for the user in his/her journey (Platzer, 2018). UPs could be considered explicit pain points; users are conscious and aware of those system issues intervening while they execute their tasks (Bruun et al., 2016). PPPs are more comprehensive since they include automatic, often unconscious, negative physiological reactions of the users. Valence-arousal PPPs are defined by high arousal and high negative emotional valence prompted by non-optimal interaction moments (Giroux-Huppé et al., 2019). Cognitive load-valence PPPs are here defined by high cognitive load and high negative valence. They are build upon the PPPs identification method (Giroux-Huppé et al., 2019) and recent research that found a negative impact of the occurrence of both high frustration and high cognitive load on the users (de Guinea et al., 2014). For this study, the 2 types of PPPs are not differentiated, they are both categorized as PPPs. Thus far, to the extent of our knowledge, no research has focused on providing guidelines on the number of subjects required for usability testing with PPPs.

1.3 Method

1.3.1 Design, Sample, and Procedure

We conducted a controlled study in the context of the implementation of a new information system in a business. The experimentation was designed to evaluate PPPs for new users in training.

Fifteen participants (3 men and 12 women) were recruited through our institution’s recruitment panel. Age ranged from 19 to 39 with a mean of 23.6 years (SD = 4.98). The user journey consisted of 5 tasks. Users had to reproduce the actions that were presented to them once in an instructional video before each task. The multi-step sequence displayed in the videos were representative of actions that new users would have to execute, such as basic data entries. As the tasks followed a linear order, the same sequence was used for every subject. The duration of the test was between 45 and 60 minutes, including the EDA baseline measures. The participants received a 20$ gift card as a compensation at the end of the experiment. They had to read and sign a consent form before beforehand. This project was approved by our institution's Institutional Review Board (IRB).

1.3.2 Apparatus

Throughout the study, participants’ facial emotions (to assess their emotional valence) and EDA (to assess their arousal) were recorded (Dirigan & Göktürk, 2011). Moment-to-moment emotional valence was measured in real-time using Facereader v6.0 (Noldus, Wageningen, Netherlands). This software detects participants' facial movements using the Facial Action Coding System (FACS) developed by Ekman and Friensen (Ekman & Friensen, 1978) to detect the 6 core emotions: happy, sad, angry, disgust, fear and surprised, with the addition of the neutral emotion (Loijens & Krips, 2018). The software computes the intensity of those emotions and scores them between 0 and 1 (Dawson, 2007). The emotional valence is then calculated as the intensity of happy (as the only positive emotion) minus the highest intensity between the negative emotions (sad, angry, scared and disgusted) (Dawson, 2007). Thus, the valence score is between 0 and 1, from negative to neutral to positive (Dawson, 2007). The Acqknowledge software (BIOPAC, Goleta, USA) was used to measure the arousal with the standardized

EDA of the participant (Dawson, 2007). A Tobii eye tracker (Tobii, Stockholm, Sweden) was used to record eye-tracking data in Tobii Studio software, which was also used to record the experiment. Event markers were also positioned manually after the test at the beginning and at the end of each task in Tobii Studio. Those event markers were used to perform the PPP analysis as well as creating a color code in Tableau® (Tableau, Seattle, USA). The Observer XT software (Noldus, Wageningen, Netherlands) and Cube HX (Montreal, Canada) was utilized to synchronize the equipment and event markers.

1.3.3 Analysis

PPPs were calculated with the statistical software SAS 9.4 using the same thresholds for valence and arousal as Giroux-Huppé et al. (Giroux-Huppé et al., 2019), i.e., moments where the subject is his ninetieth percentile for standardized EDA as well as in his tenth percentile of valence. Thus, those are moments where, compared to the user baseline state, high arousal coincides with high negative valence (Giroux-Huppé et al., 2019).

The user journey was modeled from the subtasks displayed in the instructional videos, i.e., a series of actions that the participant had to execute in order to fulfil each task. This user journey allowed to pair the PPPs with subtasks where participants experienced them by using the video replay function in Tobii Studio. These results were graphically illustrated afterward as an individual journey map with Tableau®.

1.4 Results

PPPs experienced by each participant were recorded chronologically. For each participant, we tallied the PPPs experienced, we counted new PPPs, i.e., those that were not experienced by any precedent participant, and the accumulated number of distinct PPPs we observe with each additional participant. Results show that over 15 subjects, 22 distinct PPPs were identified and 82% of those were identified after 9 participants.

Table 2. Details of psychophysiological pain points experienced per participant

Participant number	Number of PPPs	New PPPs	Cumulative PPPs	Cumulative %
1	2	2	2	9%
2	2	2	4	18%
3	5	3	7	32%
4	6	3	10	45%
5	6	4	14	64%
6	5	2	16	73%
7	0	0	16	73%
8	1	0	16	73%
9	5	2	18	82%
10	6	0	18	82%
11	3	1	19	86%
12	0	0	19	86%
13	5	2	21	95%
14	3	0	21	95%
15	5	1	22	100%

Calculations using 1000 trials of random orders were then performed to demonstrate the independence of the order (Nielsen & Landauer, 1993). The result shows that on average 80.4% (std=7.5%) of the identified PPPs were found with the first 9 participants. Those permutations are shown in the figure, in which the x-axis represents the number of participants and the y-axis represents the cumulative percentage of PPPs found.

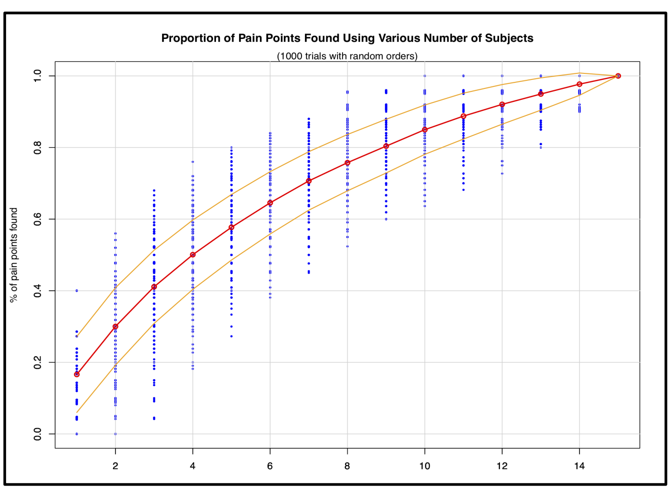

Figure 1. With 1000 permutations, the average curve (Accumulated Proportion of Detected Pain Points) follows a consonant curve to a single permutation, proving its independence.

1.5 Conclusion

As user testing with physiological measures in usability is an emerging practice, this research's objective was to provide empirically-based guidelines to UX researchers on the sample size required. Our results show that while employing psychophysiological measures in user testing with 15 participants, more than 80% of PPPs were experienced with 9 participants. Thus far, only guidelines for sample size for explicit pain points were proposed in the HCI literature (Nielsen & Landauer, 1993; Nielsen, 1994; Schmettow, 2012).

From a managerial point of view, those results can enhance usability testing planning and ROI, by providing guidelines about what could be a standard sample size to identify PPPs. It can help resolve cost management issues within organization that are transitioning to user testing with psychophysiological measures.

However, some limitations need to be considered. As this was a lab-based experiment and the instructional videos were only watched once by each participant, it did not render a totally natural experience as the participant could not just go back and watch the instructions of the action where he was blocked. Some videos were more complex and had longer duration, so they required more engagement from participants. Furthermore, as the experiment was in a specific context, external validity should be tested by investigating PPPs in other contexts.

Future research should try this approach in a two-phase process, the first one being the same as the one used in this study. The second phase should aim to evaluate if the overall number of PPPs decreased if they were addressed between development phases.

References

Albert, William et Thomas Tullis (2013). Measuring the user experience: Collecting, analyzing, and presenting usability metrics, Newnes.

Andreassi, J (2000). « Psychophysiology: Human behavior and physiological response. Lawrence erlabaum associates », *Inc., Mahwah, NJ*.

Dawson, M.E.: The Electrodermal System. Dans: Cacioppo, John T, Louis G Tassinary et Gary Berntson (2007). *Handbook of psychophysiology*, Cambridge University Press.

de Guinea, Ana Ortiz, Ryad Titah et Pierre-Majorique Leger (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.

Dirican, Ahmet Cengizhan et Mehmet Göktürk (2011). « Psychophysiological measures of human cognitive states applied in human computer interaction », *Procedia Computer Science*, vol. 3, p. 1361-1367.

Ekman, Paul et Wallace V Friesen (1978). *Manual for the facial action coding system*, Consulting Psychologists Press.

Lewis, James R (2006). « Sample sizes for usability tests: Mostly math, not magic, interactions, v. 13 n. 6 », *November+ December*.

Loijens, L et O Krips (2018). « Facereader methodology note », *A white paper by Noldus Information Technology*.

Nielsen, Jakob (1994). « Estimating the number of subjects needed for a thinking aloud test », *International journal of human-computer studies*, vol. 41, no 3, p. 385-397.

Paas, Fred, Juhani E Tuovinen, Huib Tabbers et Pascal WM Van Gerven (2003). « Cognitive load measurement as a means to advance cognitive load theory », *Educational psychologist*, vol. 38, no 1, p. 63-71.

Rohrer, Christian (2014). « When to use which user-experience research methods », *Nielsen Norman Group*.

Schmettow, Martin (2012). « Sample size in usability studies », *Commun. ACM*, vol. 55, no 4, p. 64-70.

Sweller, John, Paul Ayres et Slava Kalyuga (2011). « Measuring cognitive load », dans *Cognitive load theory*, Springer, p. 71-85.

Vernette, Eric, Marc Filser et Jean-Luc Giannelloni (2008). *Etudes marketing appliquées: De la stratégie au mix: Analyses et tests pour optimiser votre action marketing*, Dunod.

Wang, Joseph Tao-Yi. Pupil dilatation and eye tracking. Dans: Schulte-Mecklenbeck, Michael, Anton Kühberger et Joseph G Johnson (2011). *A handbook of process tracing methods for decision research: A critical review and user's guide*, Psychology Press.

Chapitre 2 : Deuxième article
The Effect of the Segmentation of Video Tutorials on User's Training Experience and Performance

Charles Lamontagne[1], Sylvain Sénécal[1], Marc Fredette[1], Bertrand Demolin[1], Shang Lin Chen[1], Romain Pourchon[2], Yohan Gaumont[2], David De Grandpré[2], Pierre-Majorique Léger[1]

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

{charles.lamontagne, sylvain.senecal, marc.fredette, shang-lin.chen, pierre-majorique.leger}@hec.ca

[2] Deloitte Digital, La Tour Deloitte, 1190, avenue des Canadiens-de-Montréal, Bureau 500, Montréal, Qc, Canada H3B 0M7

{rpourchon, ygaumont, dadegrandpre}@deloitte.ca

Abstract

This study aims to explore the effects of the segmentation of video tutorials on the implicit and explicit experience of the users in the hands-on training tasks. A laboratory experiment was conducted in the context of user training. Participants (n=20) were randomly assigned to one of two experimental conditions: with segmented or non-segmented version of the same video tutorials. The participants had to watch the video tutorials and recreate afterward the steps that were explained to them. Using self-reported and psychophysiological measures, the explicit and implicit experiences of the participants were measured. The results suggest that while the implicit experience appears to be similar across condition, the explicit experience is significantly different; segmented tutorials lead to a better self-perceived experience and a better perceived usability of the software. This study makes several contributions to the multimedia learning and the user training literature, such as the effect that the segmentation of tutorials generates on the user experience, as well as reinforcing the need to use concomitantly

complementary methods to assess the experience of the users, such as implicit and explicit measures.

Keywords : User training, User experience, psychophysiological measurements, psychophysiological pain points, change management, Video Tutorial Segmentation

2.1 Introduction

User training is an important part of change management in information system (IS) implementation (Igbaria et al., 1995). It is the first interaction of the users with the new IS, from which they evaluate if they will benefit or not of the adoption of this new IS (Beaudry & Pinsonneault, 2005). The users' evaluation process is molded by their emotions and perceptions toward the IS during this first interaction (Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015). Negative emotions, such as anger and anxiety, have a greater impact on the users' attitude toward a new IS than positive emotions, such as pleasure and enjoyment (Dohle et al., 2012; Keltner et al., 2003; Lerner & Keltner, 2000; Kim & Kankanhalli, 2009; Venkatesh et al., 2003). Therefore, it is important that the user training provide a positive experience (Beaudry & Pinsonneault, 2005; Michel et al. 2012).

Video tutorials are now widely used in user training, as they are an easy and cost-efficient method to execute user training (Van der Meij & Van Der Meij, 2013). Video tutorials are standardized as they present the same instructions in the same way to every learner, whereas trainers can adapt their teaching methods to their learners' psychological and affective states (D'Mello et al., 2006; Graesser et al., 2005; Lepper & Woolverton, 2002; Van Der Meij & Van Der Meij, 2013). Consequently, as the learners can watch these video tutorials on their own, those tutorials need to be self-sufficient for the task demonstration, as well as motivating and giving satisfaction to the learners (Tversky et al., 2002; Shea & Bidjerano, 2010). When video tutorials present more complex content, the learners can experience higher cognitive load, which can prevent them from performing all the essential cognitive processes related to learning, resulting in information loss (Mayer & Moreno, 2003).

However, this situation can be solved by the segmentation of the video tutorials, i.e., partitioning into smaller parts compared to a continuous stream of information (Spanjers et al., 2010). Shorter video tutorials demand less cognitive load as the learners have less information to process before applying the instructions right away, leading to better learning outcomes (Mayer & Moreno, 2003; Moreno, 2007, Mayer, 2005; Spanjers et al., 2010; Spanjers et al., 2012; Khacharem et al., 2013; Zacks and Tversky, 2003).

Past research in educational psychology and instructional design has studied the segmentation of video tutorials in different contexts and with different measures (Van der Meij & Van Der Meij, 2013). However, to our knowledge, there is no research that studied the effect of the segmentation of video tutorials on the user experience in the execution of the tasks for which they were trained for, i.e., in the interactions of the users with the new IS. Research suggests that the combination of complementary methods, such as explicit (self-perceived measures) and implicit (unconscious and automatic psychophysiological responses) methods, renders a deeper comprehension of the user's experience that can overcome common method bias concerns (de Guinea et al., 2014; de Guinea et al., 2012). Thus, the objective of this paper is to explore the effect of the segmentation of the video tutorials and the training tasks on the perceptions and on the psychophysiological reactions of the users. More specifically, the research question is: *To what extent segmenting video tutorials influence the user's implicit and explicit training experience?*

This article is organized as follows. Section 2.2 reviews the literature on user training with video tutorials. Section 2.3 provides the hypotheses for the study. Section 2.4 presents the study method, including participants, experimental design, procedure, stimuli, and measures. Section 2.5 displays the results, which is followed by the discussion in section 2.6.

2.2 Background

2.2.1 Change management

Much research in the IS field has been focused on explaining the drivers of user acceptance (e.g., Davis, 1989; Venkatesh et al., 2003; de Guinea et al., 2014; Darban & Polites, 2016; Mashapa et al., 2013; Legris et al., 2003; Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015)). The two most studied constructs in IS research are the perceived usefulness and the perceived ease-of-use, which are the behavioral beliefs that were at the center of the Technology Acceptance Model (TAM) (Davis, 1989). This model was developed to assist managers in the analysis of the likelihood of the implementation success of a new technology (Davis, 1989). The Unified Theory of Acceptance and Use of Technology (UTAUT) uses a similar approach as the TAM, introducing performance expectancy, social influence, and facilitating conditions to the perceived usefulness and ease of use (Venkatesh et al., 2003).

However, those models overlook emotions as key determinants toward product acceptance (de Guinea et al., 2014; Darban & Polites, 2016; Mashapa et al., 2013; Legris et al., 2003). Just as emotions play an important role in human life for prioritizing needs and organizing actions, emotions influence behavioral beliefs (Frijda et al., 2000). Research suggests that in order to fully understand the rationale behind individuals' evaluation of the usefulness and ease of use of a given IT, the emotional reactions that the IT triggers need to be included (Beaudry & Pinsonneault, 2010). Users' emotions and perceptions influence the evaluation of a new technology (Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015). This evaluation is generally conducted by the users in the first interaction with a new IT and defines the users' attitude toward the new IT, as they evaluate if it poses an opportunity or a threat to them (Beaudry & Pinsonneault, 2005). Therefore, As employees need to have a positive experience in the user training, as user training is generally the first interaction of the employees with a new IT (Beaudry & Pinsonneault, 2005; Michel et al. 2012).

2.2.2 User Training

User training is a key factor within software implementations (Gupta & Anson, 2014; Igbaria et al., 1995). The objective of user training is to train users to understand the new IT and to motivate them to use it to perform their job-related tasks (Gupta & Anson, 2014; Igbaria et al., 1995). More specifically, it aims to have learning outcomes, i.e., a change in the learner's knowledge resulting from the given instructions (Pellegrino et al., 2001). There are two main ways to measure learning outcomes, the first being the retention tests, which emphasize on the remembering, and the second being transfer tests, which emphasize on the understanding (Pellegrino et al., 2001). Executing the exact steps in a training task that were previously shown is defined as a retention test (Pellegrino et al., 2001).

The effectiveness of the user training depends on the design of training methods as well as the learning process involved (Bostrom et al., 1990). The training method consists of what is used to share the essential knowledge to the student, i.e., the medium with which the training is executed (Bostrom et al., 1990). The individuals' degree of receptiveness of information to be learned depends on their appreciation of different aspects of the learning environment, such as the instructor, the visual presentation, and the learning activities (Goleman, 1995). Research suggests that the complexity of the training tasks, which are parts of the training methods, affects learning outcomes (Maynard et al. 1997; Bolt et al., 2001; Compeau & Higgins, 1995). The complexity of a task is determined by the number of elements presented and the relations between them (Spanjers et al., 2012). Hence, providing a learning environment that is more pleasurable and comprehensible may lead to better technology adoption (Morris & Venkatesh, 2010).

The learning process is the way that the learner picks, organizes and integrates the information displayed to him with prior knowledge, which varies between people due to individual differences (Alavi et al., 2001; Bostrom et al., 1990). Individual difference constructs are generally categorized as traits or states (Bostrom et al., 1990). Traits are inherent parts of the nature of a person, such as his cognitive abilities, cognitive style (learning style), motivational traits and past experiences (Kolb, 1971; Kolb et al., 1975; Sein & Bostrom, 1989). Previous research suggests that an individual's learning style has a significant influence on the learning outcomes directly by the trainee's mental models or indirectly by the trainee's interactions with the training methods (Bohlen et al., 1997; Olfman et al., 2000).

Motivation is an important aspect of the learning process, as it stimulates the learners to spend mental effort to understand what is presented to them and to use it afterward (Davis et al., 1993; Kanfer, 1991; Pintrich & Schunk, 2002; Clark, 2015) Research suggests that learners with higher levels of motivation are prone to better learning outcomes and have higher perceived ease of use of the IS used (Davis & Bostrom., 1993; Colquitt et al., 2000). Emotions have an effect on motivation, as positive emotions tend to facilitate commitment and negative emotions tend to dissuade commitment (Bower, 1983, 1995; Ford, 1992). As humans, we are drawn to approach and keep contact with sources of positive emotions and to distance ourselves from sources of negative emotions (Bower, 1992).

2.2.3 Video Tutorials

Video tutorial is now a major instructional method in corporate software training as it is an easy and cost-efficient method to execute user training in a software implementation (Van der Meij & Van Der Meij, 2013). The rising presence of video tutorials in user training can be explained by many factors, such as the growing number of new softwares and the frequent updates of existing softwares that need to be explained to users, as well as the fact that software companies are now generally producing and distributing video tutorials rather than paper tutorials (Van der Meij & Van Der Meij, 2013; Giannakos, 2013).

The goal of a software video tutorial is to guide the users into developing procedural knowledge; procedures being sequences of steps that the users need to execute in order to achieve task completion (Plaisant and Shneiderman (2005), Bandura, 1986; Van der Meij & Van Der Meij, 2013). Video tutorials are usually in the form of a recorded demonstration, which is a demonstration-based training, i.e., an active example displaying how to achieve task completion completed by instructional voice-over explaining the actions as well as relevant software elements (Bandura, 1986; Grossman et al. 2013; Rosen et al. 2010; Van der Meij & Van Der Meij, 2013). It also needs to stimulate the learners to reflect and not simply demonstrate all the steps needed to complete the task (Van der Meij et al., 2009). Hence, it presents a scenario of use to the learners and every step to accomplish this scenario, supporting the learning and the retention of software skills (van der Meij & van der Meij, 2013).

Many studies were conducted in the field of multimedia learning, which is defined as learning from the combination of words and pictures (Mayer & Moreno, 2003). Multimedia instructions use both the users' auditory and visual working memory, which have been proven to be more efficient for users than paper tutorials, which only use the users' visual faculty to presents both the words and pictures (Moreno & Mayer, 1999; Mayer, 2005). Video instructions also provide more motivation, higher immediate post-training skills proficiency as well as superior skills retention one week after the training than paper instructions (Van der Meij & Van Der Meij, 2014).

Learners have to perform cognitive processes demanding extraneous cognitive load while listening to video tutorials, as they have to keep information in their working memory in order to link it with information that will be presented later while concomitantly process the new information that is presented to them (Leahy & Sweller, 2011; Ayres & Paas, 2007; Lowe, 1999; Mayer & Moreno, 2003). Thus, an important challenge of multimedia learning is to find the appropriate level of cognitive load, as substantial cognitive processing is needed in order to render learning outcomes, but the humans' capacity for cognitive processing is limited (Clark, 1999; Sweller, 1999; van Merriënboer, 1997; Mayer & Moreno, 2003; Catrambone & Yuasa, 2006; Linek et al., 2010; Palmiter, 1993). Cognitive overload happens when the learner's intended cognitive processing surpasses the learner's available cognitive capacity (Mayer, 2005; Mayer & Moreno, 2003). The pacing of the instructions is very important, as a slow tempo can be boring for the learners, which can lead to inattention, and a fast tempo can overload the learners, which can lead to a decrease in learning outcomes or to the learners quitting altogether (Bovair & Kieras, 1991; Linek et al., 2010). Nonessential and extraneous information can also cause cognitive overload, which implies to keep only essential information in the instructions (Mayer, 2005).

As learners can watch video tutorials alone, without trainers, those tutorials need to be self-sufficient for the task demonstration, as well as motivating and giving satisfaction to the employees (Tversky et al., 2002; Shea & Bidjerano, 2010). Trainers can adapt their training methods in response to the affective state of their learners in order to enhance learning outcomes (D'Mello et al., 2006; Graesser et al., 2005; Lepper & Woolverton, 2002), whereas video tutorials are standardized, they present the same instructions in the same way to every learner (Van Der Meij & Van Der Meij, 2013).

Thus, the design of video tutorials needs to be optimized in order to fit the learners' resources and capabilities (Kennedy, 2004; Mestre, 2012; Wouters et al., 2007). When the content of the video tutorials is more complex, it causes a high intrinsic cognitive load that can prevent the learners to perform all the essential cognitive processes, which results in a loss of information before it can be integrated with the learners' prior knowledge (Mayer & Moreno, 2003). This can be solved by the segmentation of the video tutorials, which is defined as partitioning into smaller but meaningful parts compared to a continuous stream of information (Spanjers et al., 2010). Segmentation had been linked to a beneficial effect on the learning outcomes (Mayer, 2005; Spanjers et al., 2010; Spanjers et al., 2012; Khacharem et al., 2013; Zacks and Tversky, 2003). Shorter video tutorials demand less extraneous cognitive load as it gives time for the learner to perform the cognitive activities needed and to apply the instructions right away (Mayer & Moreno, 2003; Moreno, 2007). To achieve a good segmentation, the learners need to perceive the segmented video tutorial as a standalone video tutorial with a clear beginning and ending, which can be performed by structural changes in the task or subtask completion (Zack et al., 2007; Spanjers et al., 2010).

Hence, research studied the segmentation and the ideal length of video tutorials in different contexts and with different measures and guidelines for the production of efficient instructional videos have been made from different research fields, such as educational psychology and instructional design (Van der Meij & Van Der Meij, 2013). However, to our knowledge, there is no research that studied the effect of the segmentation of video tutorials on the user experience in the hands-on training tasks, which is the first interaction of the users with the IS. User experience (UX) encompasses the whole experience of a user with a digital entity, from expectations through the interactions with the interface to the retrospective of the experience (ISO 9241-210:2010, 2.15). Hence, it concerns all aspects of the interaction of the user with a given interface, which extend to the emotional and cognitive states, thoughts, and perceptions of the user (ISO 9241-210:2010, 2.15).

2.2.4 Emotions in human-computer interactions

Research suggests that psychological states of the learner impact the learning outcomes (Olfman et al., 2005). Emotions are psychological states responding to a given stimulus (Russel, 2003; Scherer, 2005; Lazarus, 1982; Bagozzi et al., 1999). Their purpose is to organize and prioritize one's actions in order to adapt to the context and the environment (Bagozzi et al., 1999). In the context of user training, emotions are short responses to the user's interactions with the training methods and the IS (Rafaeli & Vilnai-Yavetz, 2004).

Previous research suggests that negative emotions, such as anger and anxiety, have a greater significant influence on the users' attitude toward a new IT, and subsequently on the adoption of the new IT, than positive emotions, such as pleasure and enjoyment (Dohle et al., 2012; Keltner et al., 2003; Lerner & Keltner, 2000; Kim & Kankanhalli, 2009; Venkatesh et al., 2003). More specifically, happiness has been positively connected to the intention of use (Kim & Kankanhalli, 2009) and the perceptions of ease of use of the IS (Venkatesh &

Davis, 2000). Moreover, individual's decision-making is more influenced by aroused emotional reactions (Shiv & Fedorikin, 1999).

General theories of anxiety state that frustration creates cognitive reactions that can negatively affect the expectations of a situation (Philipi et al., 1972). Frustration is defined as an emotional response to a situational opposition, i.e., an event beyond the control of the individual that restrains him/her from fulfilling his/her goal (Brown, 1954). Hence, to put this in context, a user's perceived ease of use of a given IT can be impacted if he/she experiences frustration in his interaction, as the IT does not allow the user to easily complete his/her task (de Guinea et al., 2014). Research suggests that when frustration is very high, cognitive load impacts negatively on the perceived ease of use, whereas when frustration is low, cognitive load impacts positively the perceived ease of use (de Guinea et al., 2014). Cognitive load is defined as the demands placed on working memory while executing a task (Brünken et al., 2003; Paas & Merriënboer, 1994). Cognitive load is not always disruptive; its occurrence without frustration lead to better cognitive evaluation of a given IT as adaptation and learning processes are reinforced by positive feedback (Cavanagh et al., 2010; Coyne et al., 2009). Good design also eases intrinsic cognitive load, which in turn leads to better perceived ease of use (Szabo & Kanuka, 1998).

Moreover, research suggests that ease of use diminishes users' frustration (Cox & Dale, 2002; Lazar et al., 2006). Usability of a website also reduces frustration, which inherently increases the ease of use (Pratt et al., 2004). Usability is the ability of users to use a system, product, or service to successfully perform a given task (ISO 9241-210:2010, 2.13). User acceptance of a new IT is also influenced by usability problems (UP) encountered by users within their interaction with it (Hassenzahl, 2000). Usability problems can be considered as a set of negative phenomena caused by a combination of user interface design factors and factors of usage context in the scope of the human-computer interactions (HCI) (Manakov and Ivanov, 2016). Those can be such as the user's inability to reach his/her goal, inefficient interaction and/or user's dissatisfaction (Manakov and Ivanov, 2016).

Usability testing can uncover subtle but significant UPs that can be indiscernible to designers and experts (Constantine, 2004). As a result, it provides feedback to designers on how they can improve systems, products and services (Brunn et al., 2016). Moreover, previous research has found that usability testing was useful to provide feedback to both the system customization and the user training (Kshniruk et al., 2009). Even small scale and rapid usability tests can provide valued insights that can improve the system usability (by modifications in the IS) and system learnability (by the user training) (Kshniruk et al., 2009).

2.2.5 Implicit and explicit measurements

Research studied the cognitive and emotional factors that influence the perceived ease of use and perceived usefulness (e.g., Agarwal & Karahanna, 2000; Barki et al., 2008; Karahanna & Straub, 1999). However, users' perceptions and emotions towards an IT is generally measured through self-reported measures at the end of the interaction (Bruun et al., 2016). User's experience is mostly measured with a qualitative angle with interviews and observations, with the exception of questionnaires that can include qualitative and quantitative data (Moon et al., 2016; Nenonen et al., 2018; Rawson et al., 2013). There are many explicit measurement scales that can be used to assess different aspects of user experience (UX). Those aspects are such as the usability (Brooke, 1996; Loiacono et al., 2002), the perceived workload (Hart & Staveland, 1988) or the pleasure, arousal, and dominance (Bradley & Lang, 1994). To render a complete portrayal of the experience of the users, quantitative and qualitative data must be concomitantly analyzed (Rawson et al., 2013).

A user's current experience is dependent on the physiological and psychological state of mind of the user, at that very moment (Mashapa et al., 2013). Explicit and observational measures to assess a user's experience throughout an interaction only convey a part of the user's feelings toward the digital entity (de Guinea et al., 2014). Those measures cannot assess automatic mental states that the users experience throughout their interaction that can occur without them knowing (de

Guinea & Webster, 2013; de Guinea & Markus, 2009). Qualifying the UX throughout the realization of a given task and analyzing it in synchronization with the feelings of the users allows one to understand the impact of the interface use on the user as well as to validate if the user progress toward the appropriation or the alienation of the interface (Michel et al., 2012). By assessing the self-reported emotional reactions at the end of the interaction, valuable moment-to-moment user reactions is lost (Bruun et al., 2016) and the evaluation represents more an overarching rendering that has been proven to be impacted by the Peak-End Effects (Cockburn et al., 2015; Ganglbauer et al., 2009; Hetland et al, 2018). Research suggests that the combination of complementary methods, i.e., explicit (perceptual) and implicit (psychophysiological) methods, renders a deeper comprehension of the user's experience that can overcome common method bias concerns (de Guinea et al., 2014; de Guinea et al., 2012).

Psychophysiological measures are unobtrusive and implicit manners to assess the affective and cognitive conditions of a user in HCI (de Guinea et al., 2014; Dirican and Göktürk, 2011). It is a meaningful technique to assess the moment-to-moment physiological manifestations of the psychological sentiments of the users (Andreassi, 2000). It allows users to use an IT in a realistic way and capture more reliable insights while lessening biases caused by explicit measures, as the explicit and implicit measures should be used concomitantly in order to give a better interpretation of what the users experienced (de Guinea et al., 2014; de Guinea et al., 2012; Giroux-Huppé et al., 2019). Psychophysiological measures allow assessing at any given moment emotional valence, emotional arousal, and cognitive load, which are constructs of user experience (de Guinea et al., 2012).

Many physiological measurements can be assessed, but at least 2 of those measures need to be assessed at the same time, as a sole measure can be impacted by extraneous noise (de Guinea et al., 2012; Charles & Nixon, 2019). This situation leads to the necessity of the triangulation of the physiological measures (Léger et al., 2019). Using more than one physiological measure also gives a richer comprehension of the affective and cognitive state of the user (Ganglbauer et al., 2009; Maia and Furtado, 2016). The recognition of facial expressions is a psychophysiological measure, which is used to calculate emotional valence, i.e., the emotional spectrum ranging from unpleasant (negative valence) to pleasant (positive valence) (Russell, 1980; Lane et al., 1999; Ekman and Friensen, 1978). The emotional arousal measures the excitement spectrum, which ranges from sleepiness to neutral to activation (Russell, 1980). Emotional arousal can be measured with the user's electrodermal activity (EDA) (Dawson, 2007). It quantifies the activity of the eccrine sweat glands by the electrical conductance response of the skin (Dawson, 2007). Another measure that can be assessed with psychophysiological measures is the cognitive load, which can be measured with pupillometry (Paas et al., 2003; Sweller et al., 2011). Eye-tracking technologies allow to accurately measure the pupil dilatation of the users, which can be linked to cognitive load (Wang, 2011).

The psychophysiological pain points (PPPs) identification method generates a deeper representation of the user journey as they include automatic, often unconscious, negative physiological manifestation of the user (Giroux-Huppé et al., 2019). Valence-arousal PPPs are defined as moments when the user concomitantly experiences high emotional arousal and high negative emotional valence that are caused by non-optimal interaction moments (Giroux-Huppé et al., 2019). We define cognitive load-valence PPPs as moments in which the user concomitantly experiences high cognitive load and high negative emotional valence. Cognitive load without frustration lead to better learning processes for the users (Cavanagh et al., 2010; Coyne et al., 2009), which highlights the need to identify cognitive load-valence PPPs. The PPPs identification method can be used nimbly in a change management process, as only 9 participants can be used to obtain at least 80% of PPPs (Lamontagne et al., 2019).

2.3 Hypotheses development

Learners have to perform cognitive processes demanding extraneous cognitive load while listening to video tutorials, as they have to keep information in their working memory in order to link it with information that will be presented later while concomitantly process the new information that is presented to them (Leahy & Sweller, 2011; Ayres & Paas, 2007; Lowe, 1999; Mayer & Moreno, 2003). The human capacity for cognitive processing is limited, which emphasizes the need to find the right level of cognitive load to enable learning and to prevent cognitive overload (Clark, 1999; Sweller, 1999; van Merriënboer, 1997; Mayer & Moreno, 2003; Catrambone & Yuasa, 2006; Linek et al., 2010; Palmiter, 1993). Cognitive overload happens when the learner's intended cognitive processing surpasses the learner's available cognitive capacity (Mayer, 2005; Mayer & Moreno, 2003). Cognitive overload lead to a reduction of learning outcomes as it prevents the individual to perform all the cognitive processing associated with the learning process (Mayer, 2005; Mayer & Moreno, 2003).

The complexity of a video tutorial can cause high intrinsic cognitive load, which prevents the learners to properly perform all the essential cognitive processes, resulting in learning loss before its integration with the learners' prior knowledge (Mayer & Moreno, 2003). The number of steps and the intertwining relations between those tasks determine the complexity of a task, which in turn can affect the learning outcomes (Spanjers et al., 2012; Maynard et al. 1997; Bolt et al., 2001; Compeau & Higgins, 1995). The segmentation of the video tutorials in smaller video tutorials can reduce the cognitive load (Spanjers et al., 2010). In other words, shorter video tutorials are less complex than longer video tutorials, which results in better learning outcomes as the learner can perform the essential cognitive activities needed for learning and applying the instructions right away (Mayer & Moreno, 2003; Moreno, 2007; Mayer, 2005; Spanjers et al., 2010; Spanjers et al., 2012; Khacharem et al., 2013; Zacks and Tversky, 2003). Thus, we posit:

H1: Learning a new software with segmented video tutorials results in better learning outcomes than learning the same software with non-segmented video tutorials.

Better learning outcomes should allow the user to complete all the steps required to achieve the learning tasks more easily, as they should acquire more knowledge from the instructions given (Pellegrino et al., 2001). Users can experience frustration while interacting with an IS if they perceive that the IS restrains them to complete their task (Brown, 1954). Consequently, completing the tasks more easily should reduce the frustration of the users.

Moreover, valence-arousal PPPs can be caused by non-optimal interaction moments (Giroux-Huppé et al., 2019). Those non-optimal interaction moments can be the same as usability problems, which can be such as the user's inability to reach his/her goal, inefficient interaction and/or user's dissatisfaction (Manakov and Ivanov, 2016). Hence, as segmentation will create smaller tasks, with fewer steps for the users to remember, they should encounter fewer problems and thus, experience fewer negative emotions. Thus, we posit:

H2: Learning a new software with segmented video tutorials results in fewer valence-arousal PPPs and fewer self-perceived negative emotions while putting learning into action than learning the same software with non-segmented video tutorials.

High cognitive load is not always unfavorable (Cavanagh et al., 2010; Coyne et al., 2009; de Guinea et al., 2014). The happening of cognitive load without frustration lead to better learning outcomes as well as a superior cognitive evaluation of a given IT by the users (Cavanagh et al., 2010; Coyne et al., 2009). Moreover, when frustration is low, cognitive load impacts positively the perceived ease of use, whereas when frustration is high, cognitive load impacts negatively on the perceived ease of use (de Guinea et al., 2014). The segmentation of video tutorials reduces their complexity (Mayer & Moreno, 2003). Consequently, the segmentation of training tasks reduces their complexity, as the users have to execute fewer steps to achieve task completion (Spanjers et al., 2012). Cognitive load is the demand placed on working memory while executing a task (Brünken et al., 2003; Paas & Merriënboer, 1994). We will refer to the occurrence of negative emotions (high negative valence) with high cognitive load as cognitive load-valence PPPs. Thus, we posit :

H3: Learning a new software with segmented video tutorials results in fewer cognitive load-valence PPPs and a lower self-perceived workload level while putting learning into action than learning the same software with non-segmented video tutorials.

The users' evaluation of a new IS is influenced by their emotions and perceptions toward the IS during this first interaction (Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015). Negative emotions have a greater influence on the users' attitude toward a new IT than positive emotions (Dohle et al., 2012; Keltner et al., 2003; Lerner & Keltner, 2000; Kim & Kankanhalli, 2009; Venkatesh et al., 2003). When frustration is high, cognitive load impacts negatively on the perceived ease of use, whereas when frustration is low, cognitive load impacts positively the perceived ease of use (de Guinea et al., 2014). Hence, the users' perceived ease of use of a given IT can be impacted if they experience frustration in their interactions (de Guinea et al., 2014). Users can experience frustration while interacting with an IS if they perceive that the IS restrains them to complete their task (Brown, 1954).

Moreover, segmented video tutorials are beneficial for learning outcomes (Spanjers et al., 2010; Spanjers et al. (2012); Khacharem et al., 2013; Zacks and Tversky, 2003). Consequently, the users should be able to complete all the steps required to achieve the learning tasks more easily, as they should have acquired more knowledge from the instructions given (Pellegrino et al., 2001). As a result, completing the tasks more easily should reduce the frustration of the users; frustration can be experienced when the users perceived that the IS prevent them from achieving their goal easily (Brown, 1954). This refers to usability, which is the ability of users to use a system, product, or service to perform a given task successfully (ISO 9241-210:2010, 2.13). Users' inability to reach their goal, inefficient interaction and users' dissatisfaction are some of the usability problems that can occur in HCI (Manakov and Ivanov, 2016). Usability problems encountered by the users in their interactions with an IT influence their acceptance this IT (Hassenzahl, 2000). Thus, we posit:

H4: Learning a new software with segmented video tutorials results in a higher perceived usability of the software than learning the same software with non-segmented video tutorials.

2.4 Material and Method

2.4.1 Design

A laboratory experiment was conducted to test our hypotheses. The experiment was conducted in the context of the implementation of a customer relationship management (CRM) system and it was designed to emulate user training with video tutorials. Users had to watch video tutorials explaining the sequence of actions to execute a given task and had to reproduce those sequences immediately after. A one-factor between-subject experimental design with 2 conditions was used: Segmented (S) and Not Segmented (NS).

2.4.2 Sample

This study was conducted with 20 participants 10 for each condition, who were recruited through our institution's recruitment panel. In order to emulate user training, the participants had no previous experience with the CRM tested, just like employees when they start the user training. The participants for condition S were 5 men and 5 women, ranging from 21 to 28 years old with a mean of 23.7 years (SD = 2.00). The participants for condition NS were 4 men and 6 women, ranging from 20 to 28 years old, with a mean of 23.7 years (SD = 2.75).

2.4.3 Experimental procedure

Users had to reproduce the sequences of actions that were presented to them in a video tutorial before each task. Those tasks were representative of core actions that one would have to execute as a new user, such as creating a new customer file. The training tasks were chosen in collaboration with experts in the implementation of this CRM.

The training course for this study consisted of 5 tasks, the same task sequence was used for every participant as the tasks were interconnected and followed a linear order. Those 5 tasks were the same used in Lamontagne et al. (2019). As it is presented in the next section, only 2 tasks were studied, the 3 others were filler tasks used to recreate the same learning environment for both conditions. The tutorials were presented once to the users, who did not have any control on the presentation, i.e., they could not pause or rewind the tutorial. During the execution of the tasks, a tablet mounted on a stand near the computer screen displayed the objective of the task as well as the name of every field to fill with their respective data to use.

2.4.4 Experimental stimuli

To investigate the effect of the segmentation of the video tutorials, the 2 tasks that had longer video tutorials and that were deemed as more complex as they had more steps were divided into smaller tasks. The 3 other tasks were filler tasks used to recreate the same training session for the 2 conditions. Hence, the video tutorial of Task 2 (6 minutes 04 seconds) was divided into 3 different parts and the video tutorial of Task 4 (3 minutes 57 seconds) was divided into 2 different parts. For the purpose of this paper, we will refer to Task 2 as Long Task and Task 4 as Short Task. As proposed by Mayer (2005), the separation of the tasks was done as to result into manageable but still meaningful tasks. A simple introduction and conclusion were added to each divided part in order for those to be stand-alone video tutorials, but they presented exactly the same actions to execute as the non-segmented video tutorials. This resulted in video tutorials with various durations for Long Task (1 minute 13 seconds, 2 minutes 7 seconds and 2 minutes 7 seconds) as well as for Short Task (2 minutes 18 seconds and 1 minute 43 seconds). Some research has suggested that the length of video tutorials should be in the average of 3 minutes (Chan et al., 2010) whereas other research has suggested an ideal duration between 15 and 60 seconds (Plaisant and Schneiderman, 2005). Thus, the duration of the video tutorials used in the experiment is in between those.

Our team produced the video tutorials that were presented as software training in the form of a recorded demonstration. Those videos were created with the native screen recording tool in MacOS (Apple, Cupertino, USA) with the option that shows the mouse clicks by showing black circles around the pointer. Those demonstrations displayed an order of steps and actions that the users needed to execute to achieve task completion.

The duration of the experiment was between 45 and 60 minutes, which includes the EDA baseline measures. Participants received compensation in the form of a 20$ gift card at the end of the experiment. Before the experiment, participants had to read and sign a consent form. This project was approved beforehand by our institution's Institutional Review Board (IRB).

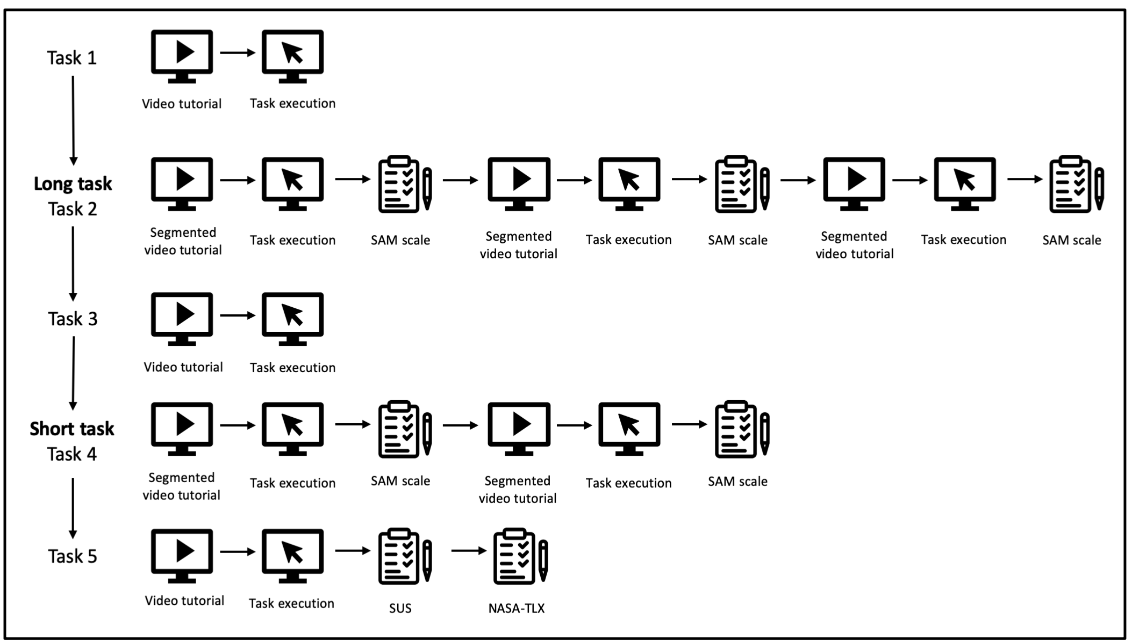

Figure 2: Experimental design for condition S

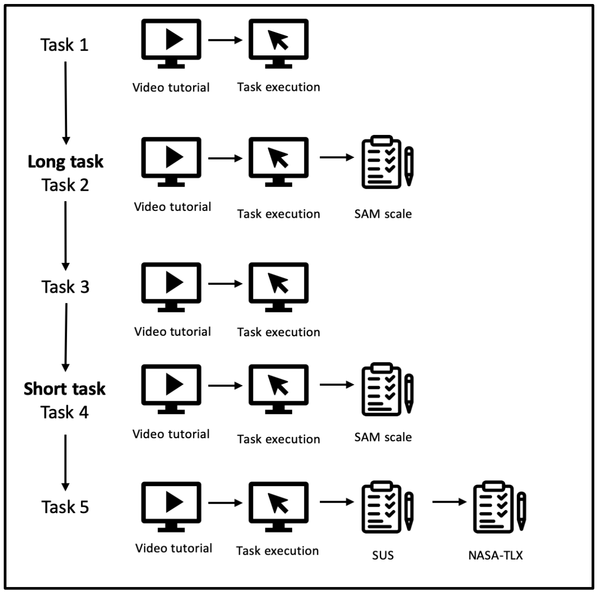

Figure 3: Experimental design for condition NS

2.4.5 Explicit Measurement

The perceived usability of the system was assessed with the System Usability Scale (SUS) at the end of the training. The SUS is a 10-item questionnaire used to evaluate the perceived usability of a system (Brooke, 1996). The even-numbered items are positively toned, and the odd-numbered items are negatively toned. Each item is divided in 5 scales steps, scaling from Strongly disagree (1) to Strongly agree (5). The participants need to complete this scale after their interactions with the system before having any debriefing or retrospective questions on their interactions (Brooke, 1996).

The emotional state was assessed with the Self-Assessment Manikin scale (SAM scale) after each training task. The SAM scale is an emotion assessment technique that uses a graphic 5-point scale to describe and measure the 3 items (pleasure, arousal, and dominance) (Bradley & Lang, 1994).

The explicit cognitive load of the participants was assessed with the raw version of the NASA Task Load Index (NASA-TLX) at the end of the training. The NASA-TLX is a multi-dimensional scale designed to obtain workload estimates. The raw version uses all the same dimensions, but it is a simplified version of the questionnaire (Hart, 2006). The NASA-TLX can be used during the completion of a task or immediately afterward (Hart & Staveland, 1988). It is the most used self-reported measurement scale to assess the cognitive load of the users (Cain, 2007). This scale is composed of six items: Mental, Physical, and Temporal Demands, Frustration, Effort, and Performance. It is useful for the assessment of the subjective workload experience. Each item is rated on a 100-point range with 5-point intervals, which are computed to give the task load index (Hart & Staveland, 1988).

2.4.6 Implicit Measurement

Three physiological measures were recorded on the participants throughout the study. Those three measures were the facial emotions of the participants, which was used to assess their emotional valence, their electrodermal activity, which was used to assess their emotional arousal, and their pupil dilation, which was used to assess their cognitive load. Facereader v6.0 (Noldus, Wageningen, Netherlands) was used to capture the emotional valence in real time. Using the Facial Action Coding System (FACS) developed by Ekman and Friensen (Ekman and Friensen, 1978), Facereader analyze participants' facial movements in order to detect the 6 core emotions : happiness, sadness, anger, disgust, fear and surprise, with the addition of the neutral emotion (Loijens and Krips, 2018). This software quantifies the intensity of those emotions, scoring them between 0 and 1 (Loijens and Krips, 2018). The emotional valence is subsequently calculated from the intensity score of the emotions as the intensity of the only positive emotion, happy, minus the highest intensity between the negative emotions, which are sad, angry, scared, and disgusted (Loijens and Krips, 2018). Consequently, the emotional valence has a value between 0 and 1, from negative to positive (Loijens and Krips, 2018).

The emotional arousal was measured with the standardized electrodermal activity of the participant with the Acqknowledge software (BIOPAC, Goleta, USA). The skin conductance response, also known as electrodermal amplitude, was used to measure the participants' excitement spectrum (Dawson, 2007). Two sensors were positioned on the hand of the user to assess this measure, which is obtained by the activity of the eccrine sweat glands by the electrical conductance response of the skin (Dawson, 2007). Sweat glands in the hands are activated by affective arousal, it is possible to convert the palmar sweating activity into the level of effective arousal (Dawson, 2007). A very low-voltage electric current is constantly passing through the two electrodes on the participant's hand, which is converted following Ohm's law. The Acqknowledge software capture and store the data, which can be used to specifically pinpoint moments of interest.

Eye-tracking data and pupils dilatation were recorded in the Tobii Studio software (Tobii, Stockholm, Sweden) with a Tobii eye tracker. This software was also used to record the experiment and to position event markers at the beginning and at the end of each video and each task. These event markers were used to measure the task completion time, which is a usability metric often used in usability tests (Sauro & Lewis, 2009). Task completion time is the time that the participants takes to complete all the steps of a task. Those event markers were also posteriorly used to analyze the PPPs. The synchronization of the apparatus and event markers was achieved by The Observer XT software (Noldus, Wageningen, Netherlands). This synchronization allowed the triangulation of the user data with Cube HX (Noldus, Wageningen, Netherlands) to obtain the PPPs (Courtemanche et al., 2019; Léger et al., 2019).

The psychophysiological measures were then used to calculate the PPPs with the statistical software SAS 9.4 (SAS, Cary, USA). The threshold used for valence-arousal PPPs was the same as Giroux-Huppé et al. (Giroux-Huppé et al., 2019), i.e., moments where the subject is his ninetieth percentile for standardized EDA as well as in his tenth percentile of valence. Therefore, those are moments where one experiences both high arousal and high negative valence compared to his own baseline state (Giroux-Huppé et al., 2019). The threshold used for the cognitive load-valence PPPs was moments where the subject is his ninetieth percentile for cognitive load as well as in his tenth percentile of valence. This threshold represents moments where the participants experienced both high cognitive load and high negative valence compared to their normal state. Then, using the video replay function in Tobii Studio, the PPPs of all participants were analyzed and labeled.

Table 3. Operationalization of Constructs

Construct	Moment of measurement	Definition	Measures
Learning performance	At the end of the training session	Participants' time to complete a task.	Tobii Studio event markers positioned at the beginning and the end of each task. *Task completion time = Time_end - Time_start*
Valence-arousal PPP (implicit)	During training tasks	Frequency of moments of negative emotions during the execution of training task, defined by the combination of low valence (lowest 10%) and high emotional arousal (highest 10%).	Valence: Facial expression recognition with the software Facereader v6.0 (Noldus, Wageningen, Netherlands) Arousal: Electrodermal activity with the Acqknowledge software (BIOPAC, Goleta, USA)
Negative emotions (explicit)	At the end of each training task.	Participants' perceptions of their emotional state during the execution of the training tasks.	Items evaluated on a 5-point scales. • Pleasure • Arousal • Dominance
Valence-cognitive load PPP (implicit)	During training tasks	Frequency of moments of negative cognitive load during the execution of training task, defined by the combination of low valence (lowest 10%) and high cognitive load (highest 10%).	Valence: Facial expression recognition with the software Facereader v6.0 (Noldus, Wageningen, Netherlands) Cognitive load: Pupil dilation with a Tobii eye-tracker and the Tobii Studio software (Tobii, Stockholm, Sweden)

Cognitive workload (explicit)	At the end of the training session	Participants' perceptions of workload of the training task and IS used.	Items evaluated on a 100-point scales with 5-point intervals • Mental demand: How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex? • Physical demand: How much physical activity was required? Was the task easy or demanding, slack or strenuous? • Temporal demand: How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid? • Performance: How successful were you in performing the task? How satisfied were you with your performance? • Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance? • Frustration: How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task? (Hart & Staveland, 1988)
Perceived usability (explicit)	At the end of the training session	Participants' perceptions of the IS used.	Items evaluated on a 5-point Likert scale. • I think that I would like to use the website frequently. • I found the website to be simple. • I thought the website was easy to use. • I think that I could use the website without the support of a technical person.

			• I found the various functions in the website were well integrated. • I thought there was a lot of consistency in the website. • I would imagine that most people would learn to use the website very quickly. • I found the website very intuitive. • I felt very confident using the website. • I could use the website without having to learn anything new. (Brooke, 1996)

2.4.7 Analysis

In order to test the difference of the implicit and explicit experiences of the participants of the 2 conditions, the results of the different measures were tested with Wilcoxon Rank Sum Tests. The 1-tailed p-values of the Wilcoxon Rank Sum Tests indicated if the differences were significant or not. The NASA-TLX and SUS results were analysed at the condition level as they were assessed at the end of the experiment. The SAM scale results, task time completion, and the 2 types of PPPs of both Long task and Short task were tested as they were assessed throughout or directly after these tasks. Only the results of Long task and Short task were tested as the 3 other tasks were the same for both conditions.

2.5 Results

All participants successfully complete all the training tasks. The task completion time for the participants of condition NS was significantly higher than the participants of condition S for both Long task ($p=.0011$) and Short task ($p=.0293$) as well as for Task 4. These results support H1.

We found no statistical difference ($p>.05$) for all types of PPPs (valence-arousal, cognitive load-valence, and the combination of the 2 types). The differences were not significant for the total number of PPPs experienced in both Long task and Short task. Thus, we found no evidence in support of H2 and H3 in our sample with the implicit measures.

The SAM scale results show that the participants of condition NS had significantly higher emotional arousal (p=.0042) as well as lower dominance (p=.0140) than those of condition S in Long task, but the emotional valence was not significantly different from condition S (p>.05). The results show no significant differences between the conditions (p>.05) for Short task. Thus, the results of the explicit measures partially support H2.

Table 4. Task completion time, PPPs and SAM scale mean results for Long task and Short task

Task	Measure	Condition S	Condition NS	P-value Difference
Long task	Task completion time	00:03:28	00:07:05	**.0011**
	Valence-arousal PPP (implicit)	.1	.3	.2911
	Cognitive load-valence PPP (implicit)	.7	1.2	.1089
	Combined PPP (implicit)	.8	1.5	.0719
	Perceived valence (explicit)	3.66	2.8	.2465
	Perceived arousal (explicit)	2.1	3.3	**.0042**
	Perceived dominance (explicit)	3.73	2.7	**.0140**
Short task	Task completion time	00:02:56	00:03:57	**.0293**
	Valence-arousal PPP (implicit)	.5	.9	.2326
	Cognitive load-valence PPP (implicit)	.2	.5	.1503
	Combined PPP (implicit)	.7	1.4	.2519
	Perceived valence (explicit)	3.3	2.9	.4837
	Perceived arousal (explicit)	2.2	2.6	.4070
	Perceived dominance (explicit)	3.45	2.9	.4369

The NASA-TLX raw score was significantly higher for the participants of condition NS (p=.0226), which suggests that the segmentation had a significant effect on the perceived workload of the tasks for the participants in retrospective, supporting H3. Moreover, we found a significant difference in the NASA-TLX subscales frustration (p=.0104) and mental demand p=.0315). We found no significant difference (p>.05) for all the other subscales.

Table 6. NASA-TLX mean scores for the 2 conditions

Item	Condition S	Condition NS	P-value
Raw TLX	18.92	31.55	**.0226**
Mental demand	33.8	53.1	**.0315**
Physical demand	3.7	4	.4211
Temporal demand	10.9	16.8	.3366
Performance	13.8	27.3	.1277
Effort	31.2	37.9	.2478
Frustration	20.1	48.4	**.0104**

The SUS mean score of the participants of condition S was significantly higher than the participants of condition NS. A higher score means that the participants of condition S perceived the usability of the IS as better than those of condition NS. This result supports H4.

Table 7. SUS mean score for the 2 conditions

SUS	Condition S	Condition NS	P-value
Mean Score	70.75	52.75	**.0476**

2.6 Discussion

The objective of this study was to explore the effect of the segmentation of the video tutorials and training task on the perceptions and on the psychophysiological reactions of the users. The results suggest that the implicit experience is the same across the 2 conditions, but the explicit experience is different. More specifically, segmented video tutorials lead to a better self-perceived experience as well as to a better perceived usability of the software used in the training.

The results show that the segmentation of the video tutorials positively impacted the task completion time, which was used to evaluate the learning performance, both in Long task and in Short task (H1). Those results are in line with recent research on segmentation, which shows that the segmentation of video tutorials has a beneficial effect on the learning outcomes (Spanjers et al., 2010; Spanjers et al. (2012); Khacharem et al., 2013; Zacks and Tversky, 2003). The number of steps explained in each segment was lower, which resulted in the reduction of the complexity of the training tasks (Spanjers et al., 2012). The results are in line with studies on the complexity of training tasks that suggest that higher complexity impacts the learning outcomes (Maynard et al. 1997; Bolt et al., 2001; Compeau & Higgins, 1995).

Additionally, the analysis of our results shows that the segmentation of video tutorials does not reduce valence-arousal PPPs (H2) and cognitive load-valence PPPs (H3) in the execution of the training task. Participants from both conditions experienced valence-arousal PPPs as well as cognitive load-valence PPPs in the execution of the training task. However, there was no significant difference between the 2 conditions for the number of PPPs experienced. This could be explained by different factors. Usability problems are caused by the combination of user interface design factors and factors of usage context (Manakov and Ivanov, 2016). As all the participants used the same IS and watched the same instructions, the context of use was the same and they were all trying to achieve the same goal. Thus, they could have encountered the same usability problems and suboptimal interaction moments, which can cause PPPs (Giroux-Huppé et al., 2019). This reinforces the need to execute user testing in order to find pain points and usability problems to provide feedback to designers on how they can improve the IS (Brunn et al., 2016). Thus, those results are in line with Kshniruk et al. (2009), user testing of the user training is useful to provide insights related to the system usability and learnability.

The segmentation of Long task impacted positively the perceived emotional arousal and the perceived dominance. The participants that watched the segmented video tutorials felt less emotionally aroused and felt more in control while executing the training task. However, they did not report a significant different emotional valence than the participants that watched the non-segmented version of the tutorials. The perceived emotional arousal, the perceived dominance, and the perceived emotional valence were not significantly different between the 2 conditions for Short task. Those results could be the effects of the reduced complexity of Long task, as it was divided into 3 segments, and of the total duration of the non-segmented version of this tutorial, which was a massive 6 minutes 7 seconds. The participants that watched the non-segmented version had to process more information and the time span between the moment where they assimilated a specific step to the moment where they applied this specific step was considerably longer. Accordingly, the results of the NASA-TLX show that the participants that watched the segmented version of the tutorials rated the perceived workload, mental demand, and frustration lower than the participants that watched the non-segmented versions of the tutorials. Hence, those findings support existing literature on video tutorials segmentation that suggest that shorter video tutorials demand less cognitive load as they present less information to process (Mayer & Moreno, 2003; Moreno, 2007).

The segmentation of the video tutorials had a positive impact on the perceived usability of the target IS. The SUS score of the participants that watched the segmented versions of the tutorials was significantly higher than the participants that watched the non-segmented versions of the tutorials (H4). This result is interesting because the proportion of valence-arousal PPPs and cognitive load-valence PPPs experienced by the participants of both conditions were not significantly different. Thus, even if they experienced the same implicit training experience, the participants that watched the non-segmented versions of the tutorials had a different perception of the usability of the IS. Additionally, the results from the NASA-TLX show that the participants that watched the non-segmented versions of the tutorials perceived a higher workload, higher mental demand and higher frustration in the training session. Hence, the explicit negative cognitive load influences the perceived ease of use, which is in line with de Guinea et al. (2014). The results of this study suggest that when frustration is very high, cognitive load impacts negatively on the perceived ease of use, whereas when frustration is low, cognitive load impacts positively the perceived ease of use (de Guinea et al., 2014). The results from the SAM scale reported that the participants that watched the non-segmented tutorial were also more emotionally aroused and felt less in control in the Long task.

There are multiple theoretical and practical implications to those results. First, results support the existing literature that suggests that the combination of explicit and implicit methods renders a deeper comprehension of the user's experience (de Guinea et al., 2014; de Guinea et al., 2012). The sole use of psychophysiological measures would have had no significant difference between the experience of the participants that watched the segmented and non-segmented versions of the video tutorials. The sole use of perceptual measures would have made some significant difference that could have been interpreted as very dissimilar training experience depending of the video tutorial version watched. Second, the Peak-End Effects literature suggests that self-reported measures assessed at the end of an experiment represents more an overarching rendering of the experience shaped by the most intense moments as well as the last moments of the experiment (Cockburn et al., 2015; Ganglbauer et al., 2009; Hetland et al, 2018). The participants in both conditions had the same version of the last task (Task 5) and they had the same implicit experience in the training session. However, the participants in both conditions had different perceptions of their training experience. This difference could be caused by many factors, but it could be connected with the fact that the users of condition NS took more time to complete both the Long and the Short task and they had more difficulty in executing the task. Third, the synthesis of the results seems to be in line with recent research that suggests that users' emotions and perceptions influence the evaluation of a new IT, which is done by the users in their first interaction with an IT (Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015). As the first interaction of the employees with a new IT is often in user training, this seems to be also in line with Morris & Venkatesh (2010); providing a more facilitating learning environment may lead to better technology adoption. In order to provide a more facilitating learning environment, research suggests that the design of video tutorials needs to be optimized in order to fit the learners' resources and capabilities (Kennedy, 2004; Mestre, 2012; Wouters et al., 2007). Our results show that the segmentation of the videos tutorials provides a better training experience and experience.

2.7 Conclusion

Our study contributes to the multimedia learning literature by evaluating the impact of the segmentation of video tutorials on the user experience, which concerns emotional and cognitive states, thoughts, and perceptions of the user (ISO 9241-210:2010, 2.15). We used explicit and implicit methods to assess the users' training experience. The combination of those complementary methods renders a deeper comprehension of the user's experience compared to the sole use of one of those measures (de Guinea et al., 2014; de Guinea et al., 2012). Past research has demonstrated that users' evaluation of new IT is influenced by factors such as emotions and perceptions (Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015). Our results show that even if the implicit experience is the same for both conditions, the explicit experience of the users that watched longer video tutorials was impacted; they consciously perceived the training tasks as more demanding and evaluated the usability of the target IS as lower than the participants of condition S. The learning outcomes of the participants of condition NS was also impacted, as their task completion time was significantly higher than the participants from condition S. Hence, our results support the video tutorials segmentation body of research (e.g.: Mayer, 2005; Spanjers et al., 2010; Spanjers et al., 2012; Khacharem et al., 2013; Zacks and Tversky, 2003; Zack et al., 2007; Spanjers et al., 2010; Mayer & Moreno, 2003; Moreno, 2007).

From a managerial point of view, as video tutorials are now vastly used in corporate software training (Van der Meij & Van Der Meij, 2013; Giannakos, 2013) and as user training is a key factor in software implementation (Gupta & Anson, 2014; Igbaria et al., 1995), our results reinforce the need to optimize video tutorials for the learners. This can be achieved by dividing the instructions into smaller segmentation in order to lessen the complexity of the training instructions and tasks. The results of our study also show that conducting user testing on the user training material can provide useful insights for the enhancement of the system usability and learnability, which was also suggested by Kshniruk et al. (2009).

However, this paper has several limitations that need to be considered for follow-up studies. First, the learning outcomes were only measured with task time completion, which only evaluate the performance in the task execution. Therefore, learning retention was not evaluated. Also, the evaluation of the user experience was only aimed on the interaction with the IS, it did not encompass the video tutorials. Thus, it would be interesting to delve into the implicit and explicit experience of the learner with the video tutorials. There is also a lack of external validity in the subjects and settings, as the participants were only students, the experiment was conducted in a controlled lab, and it was only conducted with a sole type of IS and video tutorials. Different emotions might be triggered by the context and the intrinsic motivation of the users has an effect on the learning (Yi et al., 2003). It would then be interesting to use different groups of users to compare the results, as well as different types of IS.

Moreover, the sample size used for this study was following the number of participants to find the PPPs identified by Lamontagne et al. (2019). This sample size is required to find approximately 80% of the PPPs. Hence, a follow-up study should be conducted with a larger sample size as well as in other context in order to add validity to the results.

For experimentation reason, access to the video tutorials was blocked during the execution of the task, which would not be the case in a real context. The users had no control on the video tutorials, they could not go back to parts that they did not quite understand, which would also not be the case in a real context.

References

Agarwal, Ritu et Elena Karahanna (2000). « Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage », *MIS quarterly*, p. 665-694.

Alavi, Maryam et Dorothy E Leidner (2001). « Research commentary: Technology-mediated learning—a call for greater depth and breadth of research », *Information systems research*, vol. 12, no 1, p. 1-10.

Albert, William et Thomas Tullis (2013). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*, Newnes.

Andreassi, J (2000). « Psychophysiology: Human behavior and physiological response. Lawrence erlabaum associates », *Inc., Mahwah, NJ*.

Ayres, Paul et Fred Paas (2007). « Making instructional animations more effective: A cognitive load approach », *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, vol. 21, no 6, p. 695-700.

Bagozzi, Richard P, Mahesh Gopinath et Prashanth U Nyer (1999). « The role of emotions in marketing », *Journal of the academy of marketing science*, vol. 27, no 2, p. 184-206.

Bandura, Albert (1986). « Social foundations of thought and action », *Englewood Cliffs, NJ*, vol. 1986.

Barki, Henri, Guy Paré et Claude Sicotte (2008). « Linking it implementation and acceptance via the construct of psychological ownership of information technology », *Journal of Information Technology*, vol. 23, no 4, p. 269-280.

Beaudry, Anne et Alain Pinsonneault (2005). « Understanding user responses to information technology: A coping model of user adaptation », *MIS quarterly*, vol. 29, no 3.

Beaudry, Anne et Alain Pinsonneault (2010). « The other side of acceptance: Studying the direct and indirect effects of emotions on information technology use », *MIS quarterly*, p. 689-710.

Bohlen, George A et Thomas W Ferratt (1997). « End user training: An experimental comparison of lecture versus computer-based training », *Journal of Organizational and End User Computing (JOEUC)*, vol. 9, no 3, p. 14-27.

Bolt, Melesa Altizer, Larry N Killough et Hian Chye Koh (2001). « Testing the interaction effects of task complexity in computer training using the social cognitive model », *Decision Sciences*, vol. 32, no 1, p. 1-20.

Bostrom, Robert P, Lorne Olfman et Maung K Sein (1990). « The importance of learning style in end-user training », *Mis Quarterly*, p. 101-119.

Bovair, Susan et David E Kieras (1991). « Toward a model of acquiring procedures from text », *Handbook of reading research*, vol. 2, p. 206-229.

Bower, Gordon H (1983). « Affect and cognition », *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, vol. 302, no 1110, p. 387-402.

Bower, Gordon H (1992). « How might emotions affect learning », *The handbook of emotion and memory: Research and theory*, vol. 3, p. 31.

Bradley, Margaret M et Peter J Lang (1994). « Measuring emotion: The self-assessment manikin and the semantic differential », *Journal of behavior therapy and experimental psychiatry*, vol. 25, no 1, p. 49-59.

Brooke, John (1996). « Sus-a quick and dirty usability scale », *Usability evaluation in industry*, vol. 189, no 194, p. 4-7.

Brown, James AC (1954). « The social psychology of industry ».

Brunken, Roland, Jan L Plass et Detlev Leutner (2003). « Direct measurement of cognitive load in multimedia learning », *Educational psychologist*, vol. 38, no 1, p. 53-61.

Cain, Brad (2007). *A review of the mental workload literature*, Defence Research And Development Toronto (Canada).

Catrambone, Richard et Mashiho Yuasa (2006). « Acquisition of procedures: The effects of example elaborations and active learning exercises », *Learning and Instruction*, vol. 16, no 2, p. 139-153.

Cavanagh, James F, Michael J Frank, Theresa J Klein et John JB Allen (2010). « Frontal theta links prediction errors to behavioral adaptation in reinforcement learning », *Neuroimage*, vol. 49, no 4, p. 3198-3209.

Chan, Lap Ki, Nivritti G Patil, Julie Y Chen, Jamie CM Lam, Chak S Lau et Mary SM Ip (2010). « Advantages of video trigger in problem-based learning », Medical teacher, vol. 32, no 9, p. 760-765.

Charles, Rebecca L et Jim Nixon (2019). « Measuring mental workload using physiological measures: A systematic review », *Applied ergonomics*, vol. 74, p. 221-232.

Clark, Ruth C (2011). *Developing technical training: A structured approach for developing classroom and computer-based instructional materials*, John Wiley & Sons.

Clark, Richard E (2015). « Motivating performance: Part 1—diagnosing and solving motivation problems », *Performance Improvement*, vol. 54, no 8, p. 33-43.

Colquitt, Jason A, Jeffrey A LePine et Raymond A Noe (2000). « Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research », *Journal of applied psychology*, vol. 85, no 5, p. 678.

Compeau, Deborah R et Christopher A Higgins (1995). « Application of social cognitive theory to training for computer skills », *Information systems research*, vol. 6, no 2, p. 118-143.

Constantine, Larry (2004). « Beyond user-centered design and user experience: Designing for user performance », *Cutter IT Journal*, vol. 17, no 2, p. 16-25.

Cox, Jeffrey et BG Dale (2002). « Key quality factors in web site design and use: An examination », *International Journal of Quality & Reliability Management*, vol. 19, no 7, p. 862-888.

Dawson, M.E.: The Electrodermal System. Dans: Cacioppo, John T, Louis G Tassinary et Gary Berntson (2007). *Handbook of psychophysiology*, Cambridge University Press.

D'Mello, Sidney K, Scotty D Craig, Jeremiah Sullins et Arthur C Graesser (2006). « Predicting affective states expressed through an emote-aloud procedure from autotutor's mixed-initiative dialogue », *International Journal of Artificial Intelligence in Education*, vol. 16, no 1, p. 3-28.

Darban, Mehdi et Greta L Polites (2016). « Do emotions matter in technology training? Exploring their effects on individual perceptions and willingness to learn », *Computers in Human Behavior*, vol. 62, p. 644-657.

Davis, Fred D (1989). « Perceived usefulness, perceived ease of use, and user acceptance of information technology », *MIS quarterly*, p. 319-340.

Davis, Sid A et Robert P Bostrom (1993). « Training end users: An experimental investigation of the roles of the computer interface and training methods », *MIS quarterly*, p. 61-85.

De Guinea, Ana Ortiz et M Lynne Markus (2009). « Why break the habit of a lifetime? Rethinking the roles of intention, habit, and emotion in continuing information technology use », *Mis Quarterly*, p. 433-444.

de Guinea, Ana Ortiz, Ryad Titah et Pierre-Majorique Leger (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.

De Guinea, Ana Ortiz et Jane Webster (2013). « An investigation of information systems use patterns: Technological events as triggers, the effect of time, and consequences for performance », *Mis Quarterly*, p. 1165-1188.

Dirican, Ahmet Cengizhan et Mehmet Göktürk (2011). « Psychophysiological measures of human cognitive states applied in human computer interaction », *Procedia Computer Science*, vol. 3, p. 1361-1367.

DIS, ISO (2009). « 9241-210: 2010. Ergonomics of human system interaction-part 210: Human-centred design for interactive systems », *International Standardization Organization (ISO). Switzerland*.

Dohle, Simone, Carmen Keller et Michael Siegrist (2012). « Fear and anger: Antecedents and consequences of emotional responses to mobile communication », *Journal of Risk Research*, vol. 15, no 4, p. 435-446.

Ekman, Paul et Wallace V Friesen (1978). *Manual for the facial action coding system*, Consulting Psychologists Press.

Ford, Martin E (1992). *Motivating humans: Goals, emotions, and personal agency beliefs*, Sage.

Frijda, Nico H, Antony SR Manstead et Sacha Bem (2000). *Emotions and beliefs: How feelings influence thoughts*, Cambridge University Press.

Giannakos, Michail N (2013). « Exploring the video-based learning research: A review of the literature », *British Journal of Educational Technology*, vol. 44, no 6, p. E191-E195.

Goleman, Daniel (1995). *Emotional intelligence*, New York, NY, England, Bantam Books, Inc, coll. Emotional intelligence., xiv, 352-xiv, 352 p.

Graesser, Arthur C, Natalie Person, Zhijun Lu, Moon Gee Jeon et Bethany McDaniel (2005). « Learning while holding a conversation with a computer », *Technologybased education: Bringing researchers and practitioners together*, p. 143-167.

Grossman, Rebecca, Eduardo Salas, Davin Pavlas et Michael A Rosen (2013). « Using instructional features to enhance demonstration-based training in management education », *Academy of Management Learning & Education*, vol. 12, no 2, p. 219-243.

Gupta, Saurabh et Rob Anson (2014). « Do i matter?: The impact of individual differences on a technology-mediated end user training process », *Journal of Organizational and End User Computing (JOEUC)*, vol. 26, no 2, p. 60-79.

Gupta, Saurabh, Robert P Bostrom et Mark Huber (2010). « End-user training methods: What we know, need to know », *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, vol. 41, no 4, p. 9-39.

Hart, Sandra G et Lowell E Staveland (1988). « Development of nasa-tlx (task load index): Results of empirical and theoretical research », dans *Advances in psychology*, vol 52, Elsevier, p. 139-183.

Hassenzahl, Marc (2000). « Prioritizing usability problems: Data-driven and judgement-driven severity estimates », *Behaviour & Information Technology*, vol. 19, no 1, p. 29-42.

Hetland, Audun, Joar Vittersø, Simen Oscar Bø Wie, Eirik Kjelstrup, Matthias Mittner et Tove Irene Dahl (2018). « Skiing and thinking about it: Moment-to-moment and retrospective analysis of emotions in an extreme sport », *Frontiers in psychology*, vol. 9.

Igbaria, Magid, Tor Guimaraes et Gordon B Davis (1995). « Testing the determinants of microcomputer usage via a structural equation model », *Journal of management information systems*, vol. 11, no 4, p. 87-114.

Kanfer, Ruth (1990). « Motivation theory and industrial and organizational psychology », *Handbook of industrial and organizational psychology*, vol. 1, no 2, p. 75-130.

Karahanna, Elena et Detmar W Straub (1999). « The psychological origins of perceived usefulness and ease-of-use », *Information & management*, vol. 35, no 4, p. 237-250.

Keltner, Dacher, Deborah H Gruenfeld et Cameron Anderson (2003). « Power, approach, and inhibition », *Psychological review*, vol. 110, no 2, p. 265.

Kennedy, Gregor (2004). « Promoting cognition in multimedia interactivity research », *Journal of Interactive Learning Research*, vol. 15, no 1, p. 43-61.

Khacharem, Aïmen, Ingrid AE Spanjers, Bachir Zoudji, Slava Kalyuga et Hubert Ripoll (2013). « Using segmentation to support the learning from animated soccer scenes: An effect of prior knowledge », *Psychology of Sport and Exercise*, vol. 14, no 2, p. 154-160.

Kim, Hee-Woong et Atreyi Kankanhalli (2009). « Investigating user resistance to information systems implementation: A status quo bias perspective », *MIS quarterly*, p. 567-582.

Kolb, David (1975). « Towards an applied theory of experiential learning », *Theories of Group Process.*, p. 33-56.

Kolb, David A (1971). *Individual learning styles and the learning process*, MIT.

Kushniruk, Andre W, Kristin Myers, Elizabeth M Borycki et Joseph Kannry (2009). « Exploring the relationship between training and usability: A study of the impact of usability testing on improving training and system deployment », *Stud Health Technol Inform*, vol. 143, p. 277-283.

Lane, Richard D, Phyllis ML Chua et Raymond J Dolan (1999). « Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures », *Neuropsychologia*, vol. 37, no 9, p. 989-997.

Lazar, Jonathan, Adam Jones et Ben Shneiderman (2006). « Workplace user frustration with computers: An exploratory investigation of the causes and severity », *Behaviour & Information Technology*, vol. 25, no 03, p. 239-251.

Lazarus, Richard S (1982). « Thoughts on the relations between emotion and cognition », *American psychologist*, vol. 37, no 9, p. 1019.

Leahy, Wayne et John Sweller (2011). « Cognitive load theory, modality of presentation and the transient information effect », *Applied Cognitive Psychology*, vol. 25, no 6, p. 943-951.

Léger, Pierre-Majorique, Francois Courtemanche, Marc Fredette et Sylvain Sénécal (2019). « A cloud-based lab management and analytics software for triangulated human-centered research », dans *Information systems and neuroscience*, Springer, p. 93-99.

Legris, Paul, John Ingham et Pierre Collerette (2003). « Why do people use information technology? A critical review of the technology acceptance model », *Information & management*, vol. 40, no 3, p. 191-204.

Lepper, Mark R et Maria Woolverton (2002). « The wisdom of practice: Lessons learned from the study of highly effective tutors », dans *Improving academic achievement*, Elsevier, p. 135-158.

Lerner, Jennifer S et Dacher Keltner (2000). « Beyond valence: Toward a model of emotion-specific influences on judgement and choice », *Cognition & emotion*, vol. 14, no 4, p. 473-493.

Linek, Stephanie B, Peter Gerjets et Katharina Scheiter (2010). « The speaker/gender effect: Does the speaker's gender matter when presenting auditory text in multimedia messages? », *Instructional Science*, vol. 38, no 5, p. 503-521.

Loiacono, Eleanor T, Richard T Watson et Dale L Goodhue (2002). « Webqual: A measure of website quality », *Marketing theory and applications*, vol. 13, no 3, p. 432-438.

Loijens, L et O Krips (2018). « Facereader methodology note », *A white paper by Noldus Information Technology*.

Lowe, Richard K (1999). « Extracting information from an animation during complex visual learning », *European journal of psychology of education*, vol. 14, no 2, p. 225-244.

Mayer, Richard E et Roxana Moreno (2003). « Nine ways to reduce cognitive load in multimedia learning », *Educational psychologist*, vol. 38, no 1, p. 43-52.

Mayer, Richard E et Celeste Pilegard (2005). « Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles », *The Cambridge handbook of multimedia learning*, p. 169-182.

Maynard, Douglas C et Milton D Hakel (1997). « Effects of objective and subjective task complexity on performance », *Human Performance*, vol. 10, no 4, p. 303-330.

Mestre, Lori S (2012). « Student preference for tutorial design: A usability study », *Reference Services Review*, vol. 40, no 2, p. 258-276.

Moon, Heekyung, Sung H Han, Jaemin Chun et Sang W Hong (2016). « A design process for a customer journey map: A case study on mobile services », *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 26, no 4, p. 501-514.

Moreno, Roxana (2007). « Optimising learning from animations by minimising cognitive load: Cognitive and affective consequences of signalling and segmentation methods », *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, vol. 21, no 6, p. 765-781.

Moreno, Roxana et Richard E Mayer (1999). « Cognitive principles of multimedia learning: The role of modality and contiguity », *Journal of educational psychology*, vol. 91, no 2, p. 358.

Morris, Michael G et Viswanath Venkatesh (2010). « Job characteristics and job satisfaction: Understanding the role of enterprise resource planning system implementation », *Mis Quarterly*, vol. 34, no 1.

Olfman, LORNE, Robert P Bostrom et Maung K Sein (2006). « Developing training strategies with an hci perspective », *Human-Computer Interaction and Management Information Systems: Applications*, p. 258-283.

Olfman, Lorne et Proadpran Pitsatorn (2000). « End-user training research: Status and models for the future », *Framing the domains of IT management*, p. 129-146.

Paas, Fred, Juhani E Tuovinen, Huib Tabbers et Pascal WM Van Gerven (2003). « Cognitive load measurement as a means to advance cognitive load theory », *Educational psychologist*, vol. 38, no 1, p. 63-71.

Paas, Fred GWC et Jeroen JG Van Merriënboer (1994). « Instructional control of cognitive load in the training of complex cognitive tasks », *Educational psychology review*, vol. 6, no 4, p. 351-371.

Palmiter, Susan (1993). « The effectiveness of animated demonstrations for computer-based tasks: A summary, model and future research », *Journal of Visual Languages & Computing*, vol. 4, no 1, p. 71-89.

Pellegrino, James W, Naomi Chudowsky et Robert Glaser (2001). *Knowing what students know: The science and design of educational assessment*, ERIC.

Phillips, Beeman N, Roy P Martin et Joel Meyers (1972). « Interventions in relation to anxiety in school », dans *Anxiety*, Elsevier, p. 409-468.

Pintrich, Paul R et Dale H Schunk (2002). *Motivation in education: Theory, research, and applications*, Prentice Hall.

Pratt, Jean A, Robert J Mills et Yongseog Kim (2004). « The effects of navigational orientation and user experience on user task efficiency and frustration levels », *Journal of Computer Information Systems*, vol. 44, no 4, p. 93-100.

Rafaeli, Anat et Iris Vilnai-Yavetz (2004). « Emotion as a connection of physical artifacts and organizations », *Organization Science*, vol. 15, no 6, p. 671-686.

Rawson, Alex, Ewan Duncan et Conor Jones (2013). « The truth about customer experience », *Harvard Business Review*, vol. 91, no 9, p. 90-98.

Rosen, Michael A, Eduardo Salas, Davin Pavlas, Randy Jensen, Dan Fu et Donald Lampton (2010). « Based training: A review of instructional features », *Human factors*, vol. 52, no 5, p. 596-609.

Russell, James A (1980). « A circumplex model of affect », *Journal of personality and social psychology*, vol. 39, no 6, p. 1161.

Russell, James A (2003). « Core affect and the psychological construction of emotion », *Psychological review*, vol. 110, no 1, p. 145.

Scherer, Klaus R (2005). « What are emotions? And how can they be measured? », *Social science information*, vol. 44, no 4, p. 695-729.

Sein, Maung K et Robert P Bostrom (1989). « Individual differences and conceptual models in training novice users », *Human-computer interaction*, vol. 4, no 3, p. 197-229.

Shea, Peter et Temi Bidjerano (2010). « Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments », *Computers & Education*, vol. 55, no 4, p. 1721-1731.

Shiv, Baba et Alexander Fedorikhin (1999). « Heart and mind in conflict: The interplay of affect and cognition in consumer decision making », *Journal of consumer Research*, vol. 26, no 3, p. 278-292.

Spanjers, Ingrid AE, Tamara Van Gog et Jeroen JG Van Merrienboer (2012). « Segmentation of worked examples: Effects on cognitive load and learning », *Applied Cognitive Psychology*, vol. 26, no 3, p. 352-358.

Spanjers, Ingrid AE, Pieter Wouters, Tamara Van Gog et Jeroen JG Van Merrienboer (2011). « An expertise reversal effect of segmentation in learning from animated worked-out examples », *Computers in Human Behavior*, vol. 27, no 1, p. 46-52.

Stein, Mari-Klara, Sue Newell, Erica L Wagner et Robert D Galliers (2015). « Coping with information technology: Mixed emotions, vacillation, and nonconforming use patterns », *Mis Quarterly*, vol. 39, no 2, p. 367-392.

Sweller, J (1999). « Instructional design in technical areas. Camberwell », *Victoria: ACER Press*.

Sweller, John, Paul Ayres et Slava Kalyuga (2011). « Measuring cognitive load », dans *Cognitive load theory*, Springer, p. 71-85.

Szabo, Michael et Heather Kanuka (1999). « Effects of violating screen design principles of balance, unity, and focus on recall learning, study time, and completion rates », *Journal of educational multimedia and hypermedia*, vol. 8, no 1, p. 23-42.

Tversky, Barbara, Julie Bauer Morrison et Mireille Betrancourt (2002). « Animation: Can it facilitate? », *International journal of human-computer studies*, vol. 57, no 4, p. 247-262.

Van der Meij, HANS, Joyce Karreman et Michaël Steehouder (2009). « Three decades of research and professional practice on printed software tutorials for novices », *Technical Communication*, vol. 56, no 3, p. 265-292.

van der Meij, Hans, Ilona Rensink et Jan van der Meij (2018). « Effects of practice with videos for software training », *Computers in human behavior*, vol. 89, p. 439-445.

van der Meij, Hans et Jan van der Meij (2013). « Eight guidelines for the design of instructional videos for software training », *Technical communication*, vol. 60, no 3, p. 205-228.

van der Meij, Hans et Jan Van Der Meij (2014). « A comparison of paper-based and video tutorials for software learning », *Computers & education*, vol. 78, p. 150-159.

Van Merriënboer, Jeroen JG (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*, Educational Technology.

Venkatesh, Viswanath et Fred D Davis (2000). « A theoretical extension of the technology acceptance model: Four longitudinal field studies », *Management science*, vol. 46, no 2, p. 186-204.

Venkatesh, Viswanath, Michael G Morris, Gordon B Davis et Fred D Davis (2003). « User acceptance of information technology: Toward a unified view », *MIS quarterly*, p. 425-478.

Wang, Joseph Tao-Yi. Pupil dilatation and eye tracking. Dans: Schulte-Mecklenbeck, Michael, Anton Kühberger et Joseph G Johnson (2011). *A handbook of process tracing methods for decision research: A critical review and user's guide*, Psychology Press.

Wouters, Pieter, Huib K Tabbers et Fred Paas (2007). « Interactivity in video-based models », *Educational Psychology Review*, vol. 19, no 3, p. 327-342.

Yi, Mun Y et Fred D Davis (2003). « Developing and validating an observational learning model of computer software training and skill acquisition », *Information Systems Research*, vol. 14, no 2, p. 146-169.

Zacks, Jeffrey M, Nicole K Speer, Khena M Swallow, Todd S Braver et Jeremy R Reynolds (2007). « Event perception: A mind-brain perspective », *Psychological bulletin*, vol. 133, no 2, p. 273.

Zacks, Jeffrey M et Barbara Tversky (2003). « Structuring information interfaces for procedural learning », *Journal of Experimental Psychology: Applied*, vol. 9, no 2, p. 88.

Zhang, Ping (2013). « The affective response model: A theoretical framework of affective concepts and their relationships in the ict context », *MIS quarterly*, p. 247-274.

Conclusion

Ce mémoire avait deux objectifs et a été réalisé en deux temps. Le premier objectif était de connaître le nombre d'utilisateurs minimal afin d'identifier une majorité de points de friction psychophysiologiques en test utilisateur. Le second objectif, soit l'objectif principal de ce mémoire, était d'étudier l'effet de la segmentation des tutoriels vidéos sur l'expérience vécue et l'expérience perçue des utilisateurs en contexte de formation en technologies de l'information. Plus spécifiquement, ce mémoire a permis de trouver une première ligne directrice pour le nombre d'utilisateurs nécessaire pour trouver au moins 80% des points de friction psychophysiologiques. Ce nombre a ensuite été utilisé pour étudier l'impact de la segmentation des tutoriels vidéos sur les utilisateurs. Plus précisément, ce mémoire a permis d'avoir une meilleure compréhension de l'impact de la segmentation des tutoriels vidéos sur l'expérience vécue des utilisateurs, qui a été mesurée à l'aide des points de frictions émotionnels et cognitifs, ainsi que sur l'expérience perçue des utilisateurs, qui a été mesurée par des questionnaires.

Deux expériences ont été menées pour atteindre ces objectifs. Une première expérience a été menée en laboratoire à l'hiver 2019 avec 15 participants. Ceux-ci devaient suivre la formation utilisateur qui était présentée sous la forme de 5 tutoriels vidéos et après le visionnement de chaque tutoriel, les participants devaient reproduire dans le logiciel les actions présentées. L'expérience des utilisateurs a été mesurée tout au long de l'exécution de ces actions à l'aide de mesures psychophysiologiques. La valence émotionnelle a été mesurée par la détection des émotions faciales à l'aide du logiciel Facereader, l'activation émotionnelle a été mesurée par l'activité électrodermale à l'aide du logiciel Acknowledge et la charge cognitive a été mesurée à l'aide de la pupillométrie à l'aide de Tobii Studio. Les participants ont reçu une compensation sous la forme d'une carte cadeau COOP HEC d'une valeur de 20$. Cette expérience a permis la rédaction du premier article de ce mémoire, présenté au chapitre 1.

La deuxième expérience a été conduite en laboratoire à l'été 2019 avec 2 groupes de 10 participants. Les participants devaient suivre la formation utilisateur qui leur était présentée à l'aide de tutoriels vidéos et devaient exécuter les 5 tâches présentées après le visionnement des tutoriels. Le premier groupe de participants a visionné des tutoriels vidéos non segmentés et le deuxième groupe a visionné une version segmentée des mêmes tutoriels vidéos pour 2 des 5 tâches de la formation. C'est deux tâches étaient les tâches étudiées par cette expérimentation. L'expérience vécue des participants a été mesurée au cours de l'interaction avec le logiciel à l'aide de mesures psychophysiologiques. Tout comme pour la première expérience, la valence émotionnelle a été mesurée par la détection des émotions faciales à l'aide du logiciel Facereader, l'activation émotionnelle a été mesurée par l'activité électrodermale à l'aide du logiciel Acknowledge et la charge cognitive a été mesurée à l'aide de la pupillométrie à l'aide de Tobii Studio. L'expérience perçue par les participants a été évaluée à l'aide de questionnaires. Le SAM scale a été complété après l'exécution de chacune des 2 tâches à l'étude, alors que le SUS et le NASA-TLX ont été complété à la fin de la formation. Les participants ont également reçu une compensation sous la forme d'une carte cadeau COOP HEC d'une valeur de 20$. Cette expérimentation a permis rédaction du deuxième article de ce mémoire, qui a été présenté au chapitre 2.

Ce chapitre présente un rappel des questions de recherche de ce mémoire ainsi que les principaux résultats obtenus lors de ces deux études. Puis, les contributions théoriques et les implications managériales sont exposées. Finalement, les limites et pistes de recherches futures sont abordées.

Rappel des questions de recherche

L'objectif de ce mémoire était d'étudier l'effet de la segmentation des tutoriels vidéos sur l'expérience vécue et l'expérience perçue des utilisateurs en contexte de formation utilisateur. Les questions de recherche de ce mémoire par articles étaient les suivantes :

- Article 1 : Quel est le nombre minimal d'utilisateurs nécessaires pour identifier au moins 80% des points de friction psychophysiologiques dans un parcours utilisateur ?

- Article 2 : Quel est l'impact de la segmentation des tutoriels vidéos sur l'expérience implicite et explicite des utilisateurs ?

Principaux résultats

Les résultats sont présentés selon l'ordre des questions présentées ci-dessus.

Les résultats de la première étude montrent que 82% des points de friction psychophysiologiques avaient été vécus après 9 des 15 participants. La courbe de découverte de nouveaux points de friction psychophysiologiques suivait semblablement celle de la découverte de nouveaux problèmes d'utilisabilité de Nielsen et Landauer (1993), qui atteignait un plateau après 5 utilisateurs. Après 9 utilisateurs, il est possible de découvrir de nouveaux points de friction psychophysiologiques, ceux-ci ont toutefois un moins grand retour sur investissement que pour les participants précédents.

Les résultats de la deuxième étude montrent que l'expérience vécue des utilisateurs au cours de leurs interactions avec le système d'information cible de la formation utilisateur ne diffère pas qu'ils aient visionné des tutoriels segmentés ou non segmentés. Toutefois, la segmentation des tutoriels vidéos a un impact positif sur l'expérience perçue et sur la performance des utilisateurs lors de l'exécution des instructions. Plus spécifiquement, la segmentation des tutoriels améliore l'utilisabilité perçue du système cible de la formation utilisateur et réduit la charge mentale perçue ainsi que la frustration perçue.

Contributions de l'étude

Contributions théoriques

Dans un premier temps, il y a aujourd'hui un intérêt grandissant pour l'utilisation des mesures psychophysiologiques dans les tests utilisateurs afin de fournir des expériences de qualité supérieure aux utilisateurs. Il n'y avait toutefois aucune ligne directrice qui était proposée concernant le nombre d'utilisateurs nécessaires pour réaliser des tests avec de telles mesures dans la littérature sur les interactions humain-machine. Nous contribuons à la théorie en proposant une première ligne directrice pour les tests utilisateurs avec des mesures psychophysiologiques.

Dans un deuxième temps, plusieurs recherches montrent que la formation des utilisateurs est un facteur clé pour la réussite des projets d'implantation de SI, ce qui augmente l'importance d'avoir de la formation utilisateur optimisée pour les utilisateurs (Gupta & Anson, 2014; Igbaria et al., 1995; Gupta et al., 2010; Bostrom et al., 1990). Les tutoriels vidéos sont de plus en plus utilisés pour effectuer la formation des utilisateurs lors de l'implantation de nouveaux systèmes d'information (Van der Meij & Van der Meij, 2013; Giannakos, 2013). Plusieurs recherches se sont ainsi penchées sur les avantages de la segmentation des tutoriels vidéos (voir Van der Meij & Van der Meij, 2013; Spanjers et al., 2010; Spanjers et al., 2012; Khacharem et al., 2013; Zacks & Tversky, 2003; Mayer, 2005). Les résultats obtenus dans la deuxième recherche de ce mémoire viennent appuyer ces recherches en suggérant que la segmentation a un effet positif sur l'expérience perçue des utilisateurs. La segmentation des tutoriels a un effet positif sur l'utilisabilité perçue du système cible de la formation et sur la charge mentale perçue pour l'exécution de la même formation avec des tutoriels non segmentés.

De plus, les résultats supportent la littérature existante sur l'utilité de la combinaison des mesures implicites et explicites pour avoir une compréhension plus approfondie de l'expérience des utilisateurs (de Guinea et al., 2014; de Guinea et al., 2012). Également, les résultats obtenus supportent la littérature existante sur l'influence des émotions et des perceptions des utilisateurs sur l'évaluation qu'ils portent sur une nouvelle technologie (Beaudry & Pinsonneault, 2005; Michel et al. 2012; Stein et al., 2015).

Finalement, les deux recherches montrent également qu'il est possible d'utiliser la méthode d'identification des points de friction psychophysiologique de Giroux-

Huppé et al. (2019) dans le contexte de la formation utilisateur. Auparavant, cette méthode avait été seulement testée dans le domaine du commerce électronique.

Implications pratiques

Le nombre d'utilisateurs nécessaire pour trouver les points de friction psychophysiologiques peut être utilisé en pratique. Il s'agit d'une ligne directrice permettant aux organisations de mieux se préparer et mieux planifier des tests utilisateurs avec des mesures physiologiques. Plusieurs organisations commencent à faire la transition vers ce genre de tests, cette ligne directrice pourra ainsi les guider pour assurer un bon retour sur investissement.

Les résultats obtenus montrent la segmentation des tutoriels vidéos peut améliorer l'expérience perçue des utilisateurs. Il s'agit d'un avantage de plus pour que les professionnels oeuvrant dans la production de tutoriels vidéos segmentent davantage les tutoriels vidéos utilisés lors en contexte de formation utilisateur. De plus, la réalisation de tests utilisateurs avec l'utilisation de la méthode des points de friction psychophysiologiques peut être intégrée dans les stratégies de gestion du changement afin de trouver les problèmes vécus par les employés lors de leur formation et de les rectifier.

Limites et pistes de recherches futures

Il est important de considérer quelques limites aux expérimentations qui ont été conduites dans le cadre de ce mémoire.

Tout d'abord, en ce qui concerne les participants, l'échantillon de l'étude est assez limité, puisque la majorité des participants était composée d'étudiants universitaires âgés dans la vingtaine et qui sont donc habitués d'utiliser des ordinateurs couramment. Les résultats pourraient ainsi être différents avec des personnes moins habituées avec les systèmes informatiques. Le contexte de la formation utilisateur était également spécifique, il faudrait refaire une étude en utilisant un autre contexte afin d'avoir une validité externe. Il faudrait ainsi tester avec des employés afin d'avoir une validation des résultats avec des personnes qui vivent réellement le changement vers le nouveau système, puisque ceux-ci pourraient avoir des réactions différentes et la motivation intrinsèque des utilisateurs a un effet sur l'apprentissage (Yi et al., 2003). Il faudrait également tester avec des tutoriels qui portent sur d'autres types de systèmes d'information, ce qui donnerait des niveaux de complexité différents.

Les résultats dans un contexte naturel pourraient également différer puisque les vidéos étaient présentés qu'une seule fois aux participants et que ceux-ci n'avaient aucun contrôle sur les vidéos, alors qu'en pratique ceux-ci pourraient reculer pour revenir aux endroits moins compris et faire des pauses.

Bibliographie

Agarwal, Ritu et Elena Karahanna (2000). « Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage », *MIS quarterly*, p. 665-694.

Alavi, Maryam et Dorothy E Leidner (2001). « Research commentary: Technology-mediated learning—a call for greater depth and breadth of research », *Information systems research*, vol. 12, no 1, p. 1-10.

Albert, William et Thomas Tullis (2013). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*, Newnes.

Andreassi, J (2000). « Psychophysiology: Human behavior and physiological response. Lawrence erlabaum associates », *Inc., Mahwah, NJ*.

Ayres, Paul et Fred Paas (2007). « Making instructional animations more effective: A cognitive load approach », *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, vol. 21, no 6, p. 695-700.

Bagozzi, Richard P, Mahesh Gopinath et Prashanth U Nyer (1999). « The role of emotions in marketing », *Journal of the academy of marketing science*, vol. 27, no 2, p. 184-206.

Bandura, Albert (1986). « Social foundations of thought and action », *Englewood Cliffs, NJ*, vol. 1986.

Barki, Henri, Guy Paré et Claude Sicotte (2008). « Linking it implementation and acceptance via the construct of psychological ownership of information technology », *Journal of Information Technology*, vol. 23, no 4, p. 269-280.

Beaudry, Anne et Alain Pinsonneault (2005). « Understanding user responses to information technology: A coping model of user adaptation », *MIS quarterly*, vol. 29, no 3.

Beaudry, Anne et Alain Pinsonneault (2010). « The other side of acceptance: Studying the direct and indirect effects of emotions on information technology use », *MIS quarterly*, p. 689-710.

Bohlen, George A et Thomas W Ferratt (1997). « End user training: An experimental comparison of lecture versus computer-based training », *Journal of Organizational and End User Computing (JOEUC)*, vol. 9, no 3, p. 14-27.

Bolt, Melesa Altizer, Larry N Killough et Hian Chye Koh (2001). « Testing the interaction effects of task complexity in computer training using the social cognitive model », *Decision Sciences*, vol. 32, no 1, p. 1-20.

Bostrom, Robert P, Lorne Olfman et Maung K Sein (1990). « The importance of learning style in end-user training », *Mis Quarterly*, p. 101-119.

Bovair, Susan et David E Kieras (1991). « Toward a model of acquiring procedures from text », *Handbook of reading research*, vol. 2, p. 206-229.

Bower, Gordon H (1983). « Affect and cognition », *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, vol. 302, no 1110, p. 387-402.

Bower, Gordon H (1992). « How might emotions affect learning », *The handbook of emotion and memory: Research and theory*, vol. 3, p. 31.

Bradley, Margaret M et Peter J Lang (1994). « Measuring emotion: The self-assessment manikin and the semantic differential », *Journal of behavior therapy and experimental psychiatry*, vol. 25, no 1, p. 49-59.

Brooke, John (1996). « Sus-a quick and dirty usability scale », *Usability evaluation in industry*, vol. 189, no 194, p. 4-7.

Brown, James AC (1954). « The social psychology of industry ».

Brunken, Roland, Jan L Plass et Detlev Leutner (2003). « Direct measurement of cognitive load in multimedia learning », *Educational psychologist*, vol. 38, no 1, p. 53-61.

Cain, Brad (2007). *A review of the mental workload literature*, Defence Research And Development Toronto (Canada).

Catrambone, Richard et Mashiho Yuasa (2006). « Acquisition of procedures: The effects of example elaborations and active learning exercises », *Learning and Instruction*, vol. 16, no 2, p. 139-153.

Cavanagh, James F, Michael J Frank, Theresa J Klein et John JB Allen (2010). « Frontal theta links prediction errors to behavioral adaptation in reinforcement learning », *Neuroimage*, vol. 49, no 4, p. 3198-3209.

Charles, Rebecca L et Jim Nixon (2019). « Measuring mental workload using physiological measures: A systematic review », *Applied ergonomics*, vol. 74, p. 221-232.

Chan, Lap Ki, Nivritti G Patil, Julie Y Chen, Jamie CM Lam, Chak S Lau et Mary SM Ip (2010). « Advantages of video trigger in problem-based learning », Medical teacher, vol. 32, no 9, p. 760-765.

Clark, Ruth C (2011). *Developing technical training: A structured approach for developing classroom and computer-based instructional materials*, John Wiley & Sons.

Clark, Richard E (2015). « Motivating performance: Part 1—diagnosing and solving motivation problems », *Performance Improvement*, vol. 54, no 8, p. 33-43.

Colquitt, Jason A, Jeffrey A LePine et Raymond A Noe (2000). « Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research », *Journal of applied psychology*, vol. 85, no 5, p. 678.

Compeau, Deborah R et Christopher A Higgins (1995). « Application of social cognitive theory to training for computer skills », *Information systems research*, vol. 6, no 2, p. 118-143.

Constantine, Larry (2004). « Beyond user-centered design and user experience: Designing for user performance », *Cutter IT Journal*, vol. 17, no 2, p. 16-25.

Cox, Jeffrey et BG Dale (2002). « Key quality factors in web site design and use: An examination », *International Journal of Quality & Reliability Management*, vol. 19, no 7, p. 862-888.

D'Mello, Sidney K, Scotty D Craig, Jeremiah Sullins et Arthur C Graesser (2006). « Predicting affective states expressed through an emote-aloud procedure from autotutor's mixed-initiative dialogue », *International Journal of Artificial Intelligence in Education*, vol. 16, no 1, p. 3-28.

Darban, Mehdi et Greta L Polites (2016). « Do emotions matter in technology training? Exploring their effects on individual perceptions and willingness to learn », *Computers in Human Behavior*, vol. 62, p. 644-657.

Davis, Fred D (1989). « Perceived usefulness, perceived ease of use, and user acceptance of information technology », *MIS quarterly*, p. 319-340.

Davis, Sid A et Robert P Bostrom (1993). « Training end users: An experimental investigation of the roles of the computer interface and training methods », *MIS quarterly*, p. 61-85.

Dawson, M.E.: The Electrodermal System. Dans: Cacioppo, John T, Louis G Tassinary et Gary Berntson (2007). *Handbook of psychophysiology*, Cambridge University Press.

De Guinea, Ana Ortiz et M Lynne Markus (2009). « Why break the habit of a lifetime? Rethinking the roles of intention, habit, and emotion in continuing information technology use », *Mis Quarterly*, p. 433-444.

de Guinea, Ana Ortiz, Ryad Titah et Pierre-Majorique Leger (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.

De Guinea, Ana Ortiz et Jane Webster (2013). « An investigation of information systems use patterns: Technological events as triggers, the effect of time, and consequences for performance », *Mis Quarterly*, p. 1165-1188.

Dirican, Ahmet Cengizhan et Mehmet Göktürk (2011). « Psychophysiological measures of human cognitive states applied in human computer interaction », *Procedia Computer Science*, vol. 3, p. 1361-1367.

DIS, ISO (2009). « 9241-210: 2010. Ergonomics of human system interaction-part 210: Human-centred design for interactive systems », *International Standardization Organization (ISO). Switzerland*.

Dohle, Simone, Carmen Keller et Michael Siegrist (2012). « Fear and anger: Antecedents and consequences of emotional responses to mobile communication », *Journal of Risk Research*, vol. 15, no 4, p. 435-446.

Ekman, Paul et Wallace V Friesen (1978). *Manual for the facial action coding system*, Consulting Psychologists Press.

Ford, Martin E (1992). *Motivating humans: Goals, emotions, and personal agency beliefs*, Sage.

Frijda, Nico H, Antony SR Manstead et Sacha Bem (2000). *Emotions and beliefs: How feelings influence thoughts*, Cambridge University Press.

Giannakos, Michail N (2013). « Exploring the video-based learning research: A review of the literature », *British Journal of Educational Technology*, vol. 44, no 6, p. E191-E195.

Goleman, Daniel (1995). *Emotional intelligence*, New York, NY, England, Bantam Books, Inc, coll. Emotional intelligence., xiv, 352-xiv, 352 p.

Graesser, Arthur C, Natalie Person, Zhijun Lu, Moon Gee Jeon et Bethany McDaniel (2005). « Learning while holding a conversation with a computer », *Technologybased education: Bringing researchers and practitioners together*, p. 143-167.

Grossman, Rebecca, Eduardo Salas, Davin Pavlas et Michael A Rosen (2013). « Using instructional features to enhance demonstration-based training in management education », *Academy of Management Learning & Education*, vol. 12, no 2, p. 219-243.

Gupta, Saurabh et Rob Anson (2014). « Do i matter?: The impact of individual differences on a technology-mediated end user training process », *Journal of Organizational and End User Computing (JOEUC)*, vol. 26, no 2, p. 60-79.

Gupta, Saurabh, Robert P Bostrom et Mark Huber (2010). « End-user training methods: What we know, need to know », *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, vol. 41, no 4, p. 9-39.

Hart, Sandra G et Lowell E Staveland (1988). « Development of nasa-tlx (task load index): Results of empirical and theoretical research », dans *Advances in psychology*, vol 52, Elsevier, p. 139-183.

Hassenzahl, Marc (2000). « Prioritizing usability problems: Data-driven and judgement-driven severity estimates », *Behaviour & Information Technology*, vol. 19, no 1, p. 29-42.

Hetland, Audun, Joar Vittersø, Simen Oscar Bø Wie, Eirik Kjelstrup, Matthias Mittner et Tove Irene Dahl (2018). « Skiing and thinking about it: Moment-to-moment and retrospective analysis of emotions in an extreme sport », *Frontiers in psychology*, vol. 9.

Igbaria, Magid, Tor Guimaraes et Gordon B Davis (1995). « Testing the determinants of microcomputer usage via a structural equation model », *Journal of management information systems*, vol. 11, no 4, p. 87-114.

Kanfer, Ruth (1990). « Motivation theory and industrial and organizational psychology », *Handbook of industrial and organizational psychology*, vol. 1, no 2, p. 75-130.

Karahanna, Elena et Detmar W Straub (1999). « The psychological origins of perceived usefulness and ease-of-use », *Information & management*, vol. 35, no 4, p. 237-250.

Keltner, Dacher, Deborah H Gruenfeld et Cameron Anderson (2003). « Power, approach, and inhibition », *Psychological review*, vol. 110, no 2, p. 265.

Kennedy, Gregor (2004). « Promoting cognition in multimedia interactivity research », *Journal of Interactive Learning Research*, vol. 15, no 1, p. 43-61.

Khacharem, Aïmen, Ingrid AE Spanjers, Bachir Zoudji, Slava Kalyuga et Hubert Ripoll (2013). « Using segmentation to support the learning from animated soccer scenes: An effect of prior knowledge », *Psychology of Sport and Exercise*, vol. 14, no 2, p. 154-160.

Kim, Hee-Woong et Atreyi Kankanhalli (2009). « Investigating user resistance to information systems implementation: A status quo bias perspective », *MIS quarterly*, p. 567-582.

Kolb, David (1975). « Towards an applied theory of experiential learning », *Theories of Group Process.*, p. 33-56.

Kolb, David A (1971). *Individual learning styles and the learning process*, MIT.

Kushniruk, Andre W, Kristin Myers, Elizabeth M Borycki et Joseph Kannry (2009). « Exploring the relationship between training and usability: A study of the impact of usability testing on improving training and system deployment », *Stud Health Technol Inform*, vol. 143, p. 277-283.

Lane, Richard D, Phyllis ML Chua et Raymond J Dolan (1999). « Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures », *Neuropsychologia*, vol. 37, no 9, p. 989-997.

Lazar, Jonathan, Adam Jones et Ben Shneiderman (2006). « Workplace user frustration with computers: An exploratory investigation of the causes and severity », *Behaviour & Information Technology*, vol. 25, no 03, p. 239-251.

Lazarus, Richard S (1982). « Thoughts on the relations between emotion and cognition », *American psychologist*, vol. 37, no 9, p. 1019.

Leahy, Wayne et John Sweller (2011). « Cognitive load theory, modality of presentation and the transient information effect », *Applied Cognitive Psychology*, vol. 25, no 6, p. 943-951.

Léger, Pierre-Majorique, Francois Courtemanche, Marc Fredette et Sylvain Sénécal (2019). « A cloud-based lab management and analytics software for triangulated human-centered research », dans *Information systems and neuroscience*, Springer, p. 93-99.

Legris, Paul, John Ingham et Pierre Collerette (2003). « Why do people use information technology? A critical review of the technology acceptance model », *Information & management*, vol. 40, no 3, p. 191-204.

Lepper, Mark R et Maria Woolverton (2002). « The wisdom of practice: Lessons learned from the study of highly effective tutors », dans *Improving academic achievement*, Elsevier, p. 135-158.

Lerner, Jennifer S et Dacher Keltner (2000). « Beyond valence: Toward a model of emotion-specific influences on judgement and choice », *Cognition & emotion*, vol. 14, no 4, p. 473-493.

Lewis, James R (2006). « Sample sizes for usability tests: Mostly math, not magic, interactions, v. 13 n. 6 », *November+ December*.

Linek, Stephanie B, Peter Gerjets et Katharina Scheiter (2010). « The speaker/gender effect: Does the speaker's gender matter when presenting auditory text in multimedia messages? », *Instructional Science*, vol. 38, no 5, p. 503-521.

Loiacono, Eleanor T, Richard T Watson et Dale L Goodhue (2002). « Webqual: A measure of website quality », *Marketing theory and applications*, vol. 13, no 3, p. 432-438.

Loijens, L et O Krips (2018). « Facereader methodology note », *A white paper by Noldus Information Technology*.

Lowe, Richard K (1999). « Extracting information from an animation during complex visual learning », *European journal of psychology of education*, vol. 14, no 2, p. 225-244.

Mayer, Richard E et Roxana Moreno (2003). « Nine ways to reduce cognitive load in multimedia learning », *Educational psychologist*, vol. 38, no 1, p. 43-52.

Mayer, Richard E et Celeste Pilegard (2005). « Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles », *The Cambridge handbook of multimedia learning*, p. 169-182.

Maynard, Douglas C et Milton D Hakel (1997). « Effects of objective and subjective task complexity on performance », *Human Performance*, vol. 10, no 4, p. 303-330.

Mestre, Lori S (2012). « Student preference for tutorial design: A usability study », *Reference Services Review*, vol. 40, no 2, p. 258-276.

Moon, Heekyung, Sung H Han, Jaemin Chun et Sang W Hong (2016). « A design process for a customer journey map: A case study on mobile services », *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 26, no 4, p. 501-514.

Moreno, Roxana (2007). « Optimising learning from animations by minimising cognitive load: Cognitive and affective consequences of signalling and segmentation methods », *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, vol. 21, no 6, p. 765-781.

Moreno, Roxana et Richard E Mayer (1999). « Cognitive principles of multimedia learning: The role of modality and contiguity », *Journal of educational psychology*, vol. 91, no 2, p. 358.

Morris, Michael G et Viswanath Venkatesh (2010). « Job characteristics and job satisfaction: Understanding the role of enterprise resource planning system implementation », *Mis Quarterly*, vol. 34, no 1.

Nielsen, Jakob (1994). « Estimating the number of subjects needed for a thinking aloud test », *International journal of human-computer studies*, vol. 41, no 3, p. 385-397.

Olfman, LORNE, Robert P Bostrom et Maung K Sein (2006). « Developing training strategies with an hci perspective », *Human-Computer Interaction and Management Information Systems: Applications*, p. 258-283.

Olfman, Lorne et Proadpran Pitsatorn (2000). « End-user training research: Status and models for the future », *Framing the domains of IT management*, p. 129-146.

Paas, Fred, Juhani E Tuovinen, Huib Tabbers et Pascal WM Van Gerven (2003). « Cognitive load measurement as a means to advance cognitive load theory », *Educational psychologist*, vol. 38, no 1, p. 63-71.

Paas, Fred GWC et Jeroen JG Van Merriënboer (1994). « Instructional control of cognitive load in the training of complex cognitive tasks », *Educational psychology review*, vol. 6, no 4, p. 351-371.

Palmiter, Susan (1993). « The effectiveness of animated demonstrations for computer-based tasks: A summary, model and future research », *Journal of Visual Languages & Computing*, vol. 4, no 1, p. 71-89.

Pellegrino, James W, Naomi Chudowsky et Robert Glaser (2001). *Knowing what students know: The science and design of educational assessment*, ERIC.

Phillips, Beeman N, Roy P Martin et Joel Meyers (1972). « Interventions in relation to anxiety in school », dans *Anxiety*, Elsevier, p. 409-468.

Pintrich, Paul R et Dale H Schunk (2002). *Motivation in education: Theory, research, and applications*, Prentice Hall.

Pratt, Jean A, Robert J Mills et Yongseog Kim (2004). « The effects of navigational orientation and user experience on user task efficiency and frustration levels », *Journal of Computer Information Systems*, vol. 44, no 4, p. 93-100.

Rafaeli, Anat et Iris Vilnai-Yavetz (2004). « Emotion as a connection of physical artifacts and organizations », *Organization Science*, vol. 15, no 6, p. 671-686.

Rawson, Alex, Ewan Duncan et Conor Jones (2013). « The truth about customer experience », *Harvard Business Review*, vol. 91, no 9, p. 90-98.

Rohrer, Christian (2014). « When to use which user-experience research methods », *Nielsen Norman Group*.

Rosen, Michael A, Eduardo Salas, Davin Pavlas, Randy Jensen, Dan Fu et Donald Lampton (2010). « Based training: A review of instructional features », *Human factors*, vol. 52, no 5, p. 596-609.

Russell, James A (1980). « A circumplex model of affect », *Journal of personality and social psychology*, vol. 39, no 6, p. 1161.

Russell, James A (2003). « Core affect and the psychological construction of emotion », *Psychological review*, vol. 110, no 1, p. 145.

Scherer, Klaus R (2005). « What are emotions? And how can they be measured? », *Social science information*, vol. 44, no 4, p. 695-729.

Schmettow, Martin (2012). « Sample size in usability studies », *Commun. ACM*, vol. 55, no 4, p. 64-70.

Schulte-Mecklenbeck, Michael, Anton Kühberger et Joseph G Johnson (2011). *A handbook of process tracing methods for decision research: A critical review and user's guide*, Psychology Press.

Sein, Maung K et Robert P Bostrom (1989). « Individual differences and conceptual models in training novice users », *Human-computer interaction*, vol. 4, no 3, p. 197-229.

Shea, Peter et Temi Bidjerano (2010). « Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments », *Computers & Education*, vol. 55, no 4, p. 1721-1731.

Shiv, Baba et Alexander Fedorikhin (1999). « Heart and mind in conflict: The interplay of affect and cognition in consumer decision making », *Journal of consumer Research*, vol. 26, no 3, p. 278-292.

Spanjers, Ingrid AE, Tamara Van Gog et Jeroen JG Van Merrienboer (2012). « Segmentation of worked examples: Effects on cognitive load and learning », *Applied Cognitive Psychology*, vol. 26, no 3, p. 352-358.

Spanjers, Ingrid AE, Pieter Wouters, Tamara Van Gog et Jeroen JG Van Merrienboer (2011). « An expertise reversal effect of segmentation in learning from animated worked-out examples », *Computers in Human Behavior*, vol. 27, no 1, p. 46-52.

Stein, Mari-Klara, Sue Newell, Erica L Wagner et Robert D Galliers (2015). « Coping with information technology: Mixed emotions, vacillation, and nonconforming use patterns », *Mis Quarterly*, vol. 39, no 2, p. 367-392.

Sweller, J (1999). « Instructional design in technical areas. Camberwell », *Victoria: ACER Press*.

Sweller, John, Paul Ayres et Slava Kalyuga (2011). « Measuring cognitive load », dans *Cognitive load theory*, Springer, p. 71-85.

Szabo, Michael et Heather Kanuka (1999). « Effects of violating screen design principles of balance, unity, and focus on recall learning, study time, and completion rates », *Journal of educational multimedia and hypermedia*, vol. 8, no 1, p. 23-42.

Tversky, Barbara, Julie Bauer Morrison et Mireille Betrancourt (2002). « Animation: Can it facilitate? », *International journal of human-computer studies*, vol. 57, no 4, p. 247-262.

Van der Meij, HANS, Joyce Karreman et Michaël Steehouder (2009). « Three decades of research and professional practice on printed software tutorials for novices », *Technical Communication*, vol. 56, no 3, p. 265-292.

van der Meij, Hans, Ilona Rensink et Jan van der Meij (2018). « Effects of practice with videos for software training », *Computers in human behavior*, vol. 89, p. 439-445.

van der Meij, Hans et Jan van der Meij (2013). « Eight guidelines for the design of instructional videos for software training », *Technical communication*, vol. 60, no 3, p. 205-228.

van der Meij, Hans et Jan Van Der Meij (2014). « A comparison of paper-based and video tutorials for software learning », *Computers & education*, vol. 78, p. 150-159.

Van Merriënboer, Jeroen JG (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*, Educational Technology.

Venkatesh, Viswanath et Fred D Davis (2000). « A theoretical extension of the technology acceptance model: Four longitudinal field studies », *Management science*, vol. 46, no 2, p. 186-204.

Venkatesh, Viswanath, Michael G Morris, Gordon B Davis et Fred D Davis (2003). « User acceptance of information technology: Toward a unified view », *MIS quarterly*, p. 425-478.

Vernette, Eric, Marc Filser et Jean-Luc Giannelloni (2008). *Etudes marketing appliquées: De la stratégie au mix: Analyses et tests pour optimiser votre action marketing*, Dunod.

Wang, Joseph Tao-Yi. Pupil dilatation and eye tracking. Dans: Schulte-Mecklenbeck, Michael, Anton Kühberger et Joseph G Johnson (2011). *A handbook of process tracing methods for decision research: A critical review and user's guide*, Psychology Press.

Wouters, Pieter, Huib K Tabbers et Fred Paas (2007). « Interactivity in video-based models », *Educational Psychology Review*, vol. 19, no 3, p. 327-342.

Yi, Mun Y et Fred D Davis (2003). « Developing and validating an observational learning model of computer software training and skill acquisition », *Information Systems Research*, vol. 14, no 2, p. 146-169.

Zacks, Jeffrey M, Nicole K Speer, Khena M Swallow, Todd S Braver et Jeremy R Reynolds (2007). « Event perception: A mind-brain perspective », *Psychological bulletin*, vol. 133, no 2, p. 273.

Zacks, Jeffrey M et Barbara Tversky (2003). « Structuring information interfaces for procedural learning », *Journal of Experimental Psychology: Applied*, vol. 9, no 2, p. 88.

Zhang, Ping (2013). « The affective response model: A theoretical framework of affective concepts and their relationships in the ict context », *MIS quarterly*, p. 247-274.

Annexes

Annexe 1 - Table des participants du premier article

Participant	Sexe	Âge
p10	Femme	19
p11	Homme	21
p12	Femme	19
p34	Homme	23
p35	Femme	19
p36	Femme	24
p37	Femme	23
p38	Femme	22
p39	Femme	21
p40	Femme	39
p41	Femme	24
p42	Femme	24
p43	Femme	24
p44	Homme	23
p45	Femme	29

Annexe 2 - Table des participants du deuxième article

Participant	Sexe	Âge	Condition
p03	Homme	28	S
p04	Femme	24	S
p05	Femme	21	S
p06	Femme	23	S
p07	Homme	24	S
p08	Femme	21	S
p09	Homme	24	S
p10	Homme	25	S
p11	Femme	23	S
p12	Homme	24	S
p14	Femme	24	NS
p15	Homme	27	NS
p17	Homme	21	NS
p18	Femme	22	NS
p19	Femme	23	NS
p20	Homme	28	NS
p22	Homme	20	NS
p23	Femme	23	NS
p24	Femme	27	NS
p25	Femme	22	NS