[Inner endpaper]

HEC MONTRÉAL
École affiliée à l'Université de Montréal

Two Essays on the Use of Cognitive Load in Information Systems Design

par
Seyedmohammadmahdi Mirhoseini

Thèse présentée en vue de l'obtention du grade de Ph. D. en administration
(option Information technologies)

Août 2018

HEC MONTRÉAL
École affiliée à l'Université de Montréal

Cette thèse intitulée :

Two Essays on the Use of Cognitive Load in Information Systems Design

Présentée par :

Seyedmohammadmahdi Mirhoseini

a été évaluée par un jury composé des personnes suivantes :

Renaud Legoux
HEC Montréal
Président-rapporteur

Pierre-Majorique Léger
HEC Montréal
Directeur de recherche

Sylvain Sénécal
HEC Montréal
Codirecteur de recherche

Anne Beaudry
Concordia University
Membre du jury

René Riedl
University of Linz
Examinateur externe

Résumé

Bien que le concept de charge cognitive imposée aux utilisateurs ait été étudié dans le domaine des systèmes d'information pendant de nombreuses années, certaines difficultés de mesure ont empêché les chercheurs d'en explorer différents aspects. L'approche « NeuroIS », utilisant les outils et les théories des neurosciences dans le domaine des SI, nous fournit un outil puissant pour étudier les processus cognitifs des utilisateurs. En utilisant l'électroencéphalographie (EEG), les chercheurs sont en mesure de mesurer la charge cognitive avec une précision et une résolution temporelle élevées, et celle-ci peut être étudiée dans divers contextes d'interaction entre utilisateurs et TI, tels que les achats en ligne.

Cette thèse, composée de deux essais, présente différentes mesures du concept de charge cognitive et montrent comment ces dernières peuvent être utilisées pour évaluer un artefact TI. Le premier essai comprend deux expériences dont l'objectif est de mesurer la charge cognitive puis de tester ces mesures en pratique. La première expérience détermine la charge de travail instantanée de l'utilisateur et en extrait trois mesures objectives: charge accumulée, charge moyenne et charge de pointe. Cette expérience explore également l'effet de la difficulté et de l'incertitude de la tâche sur les mesures objectives et subjectives de la charge de travail. Dans la seconde expérience, nous utilisons le concept de charge accumulée pour évaluer la commodité d'un site Web d'achat en ligne et nous montrons comment cette charge accumulée est associée à la satisfaction de l'expérience d'achat. Finalement, le deuxième essai examine l'effet combiné du tri des produits et de l'objectif des utilisateurs sur leur charge de travail, en utilisant plusieurs mesures de la charge cognitive. Cette thèse contribue au champ des SI en introduisant de nouvelles mesures du construit de charge de travail et les établit comme critères pour évaluer les artefacts TI.

Mots clés : Charge cognitive, NeuroIS, Conception des SI, Charge accumulée, EEG, PGI

Méthodes de recherche : Expérience en laboratoire

Abstract

While user cognitive load has been studied in the field of information systems for many years, measurement difficulties have prevented researchers from exploring different aspects of it. NeuroIS, which seeks to use neuroscience tools and theories in the IS field, provides us with a powerful tool to study users' cognitive processes. Using Electroencephalography (EEG), researchers are able to measure cognitive load with high precision and temporal resolution, and it can be studied in various user-IT interaction contexts such as online shopping

This thesis, based on two essays, aims to introduce different measures of the cognitive load construct, and to show how they can be used to evaluate an IT artifact. The first essay includes two experiments to address both the challenge of measuring cognitive load and also testing new measures in practice. The first experiment measures the user's instantaneous cognitive load and extracts three objective measures from it: accumulated load, average load, and peak load. Then, the effect of task difficulty and task uncertainty on both objective and subjective measures of cognitive load is explored. In the second experiment of this essay we use the accumulated load construct to evaluate the convenience of an online shopping website, and show how accumulated load is associated with satisfaction with the shopping experience. Finally, the second essay examines the contingent effect of product sorting and the users goal on user workload using multiple measures of cognitive load. This thesis contributes to the IS field by introducing new measures of the workload construct and establishing them as criteria for evaluating IT artifacts.

Keywords : Cognitive Load, NeuroIS, IS design, Accumulated load, EEG, ERP

Research methods : Laboratory Experiment

Table of contents

List of Figures

List of Tables

List of acronyms

IT	Information Technology
IS	Information Systems
NeuroIS	Neuro-Information Systems
EEG	Electroencephalography
ERP	Event Related Potentials

To Niloufar,

Acknowledgements

This thesis would not have been completed without the help of many individuals. First and foremost, I would like to thank my supervisors Pierre-Majorique Léger and Sylvain Sénécal for guiding me throughout this journey and showing me the right path.

I am thankful of my family members for always being beside me during ups and downs of my Ph.D. I thank Niloufar Afsariardchi, my eternal friend and companion, who always encouraged me to accomplish this difficult challenge. Maman, Baba, Taha, and Maryam, I am sincerely blessed for having you in my life, without your love and support I would not be who I am today.

I would like to thank my friends in the Ph.D. program Thibault Coulon, Reza Ghaffari, and Hoda Nasseri who had been a great help to me. My special thanks to Omid Zobeiri, Alborz Rezazadeh, and Shahab Bakhtiari who greatly helped me with the neuroscience aspect of this interdisciplinary work.

Last but not Least, I would like to express my deepest gratitude to all professors of the IT department at HEC Montreal and Montreal's joint Ph.D. program who taught me how to conduct research. My special thanks to professors Suzanne Rivard, Ann Beaudry, and Ryad Titah who showed me how to enjoy doing academic research.

Preface

All articles of this thesis are co-authored by Seyedmohammadmahdi Mirhoseini, Pierre-Majorique Léger, and Sylvain Sénécal.

Chapter 1- Thesis Introduction

Cognitive load (or mental workload) can be defined as the amount of working memory resources allocated to perform a task (DeStefano & LeFevre, 2007). The cognitive load construct has been studied in information systems (IS) research as a measure — mostly a self-perceived measure— of efficiency, and has been used as an element of decision quality to evaluate how decision support systems can help users reduce their level of cognitive load (Todd & Benbasat, 1992, Todd & Benbasat, 1999). In recent years, researchers have started to apply this construct in design science research (Schmutz et al., 2010, Gwizdka, 2010, lo Storto, 2013). IS researchers are mostly interested in measuring the user's cognitive load during human-IT interaction and examining the effect of IT-induced cognitive load on the user's behavior.

The amount of load imposed by IT artifacts on a user's working memory is an important criterion for evaluating efficiency because people's cognitive resources are limited in nature (Wickens, 2002). Efficient use of these resources prevents users from being overloaded by the system demands thus experiencing negative consequences (lo Storto, 2013). However, measurement difficulties have prevented IS researchers from thoroughly investigating this construct (Paas & Sweller, 2012, Schmutz et al., 2010). Traditional self-perceived measures of cognitive load are subject to different types of bias (Dimoka et al., 2011) and unable to capture all variations of the cognitive load construct. De Guinea et al., (2014) distinguish between the explicit (self-perceived) and implicit (automatic or unconscious antecedents of cognitive beliefs since there are mental processes beyond a user's consciousness which nevertheless affect his/her behavior.

Recent developments in the NeuroIS field and the use of neuroscience tools and theories have enabled IS researchers to uncover different aspects of cognitive load. Besides their ability to capture automatic and unconscious mental processes, tools such as Electroencephalography (EEG) increase the temporal resolution of measuring cognitive load (Riedl et al., 2014), providing us with a richer measurement of the construct. Cognitive load can be measured over any time period during user-IT interaction. For instance, user cognitive load while working with a specific menu in a software application can be captured using EEG. This capability provides IS researchers with a powerful tool to inform IT design.

Information systems design research includes building and evaluating IT artifacts (Hevner, 2007). To improve existing artifact design, researchers must evaluate IT artifacts against an identified business need. Neuroscience tools and theories can contribute to this cycle by identifying the neurophysiological determinants of such business needs (Brocke et al., 2013) which can then be employed as criteria for design.

Prior to using the concept of working memory load in an information systems context, we need to identify its components and explain the processes within working memory. A good conceptual model of working memory is important in helping IS researchers understand how it functions. Research in the fields of psychology and neuroscience has employed various approaches to theorizing working memory. Differences between them derive chiefly from the methods used, and also the type of phenomena they emphasize. In this thesis, Baddeley's theory of working memory (Baddeley, 2007) is used since it is well established in the field of psychology and has also been strongly confirmed by neuroimaging experiments. The latter is of great importance since this study also measures cognitive load using a neuroimaging tool (EEG).

Working memory theories have evolved through the years and have been tested against competing theories. The preliminary conception of brain function was a two-component model consisting of a short term memory (STM) and a long term memory (LTM). This line of theories proposed that short-term memory is simply a temporary memory storage, and information can be transferred to long-term memory upon repetition and through learning practices. Although some evidence was found in the early years to support this model, further experiments rejected the idea that short term memory is simply a storage system. The term "working memory" was then used to emphasize the functional role of this system in the brain. Experiments showed that working memory is responsible for a wide range of storage and processing tasks; thus it is very likely that the system comprises multiple components (Smith, Jonides, & Koeppe, 1996).

It is now well established that there are separate storage systems with different responsibilities within working memory, along with a processor component which performs critical functions such as coordinating storage systems and managing attention (Baddeley, 2007). Figure 1 shows the four-component model proposed by (Baddeley, 2007) which is composed of three storage systems: visuospatial sketchpad, phonological loop, and episodic buffer which are coordinated by the fourth component, central executive.

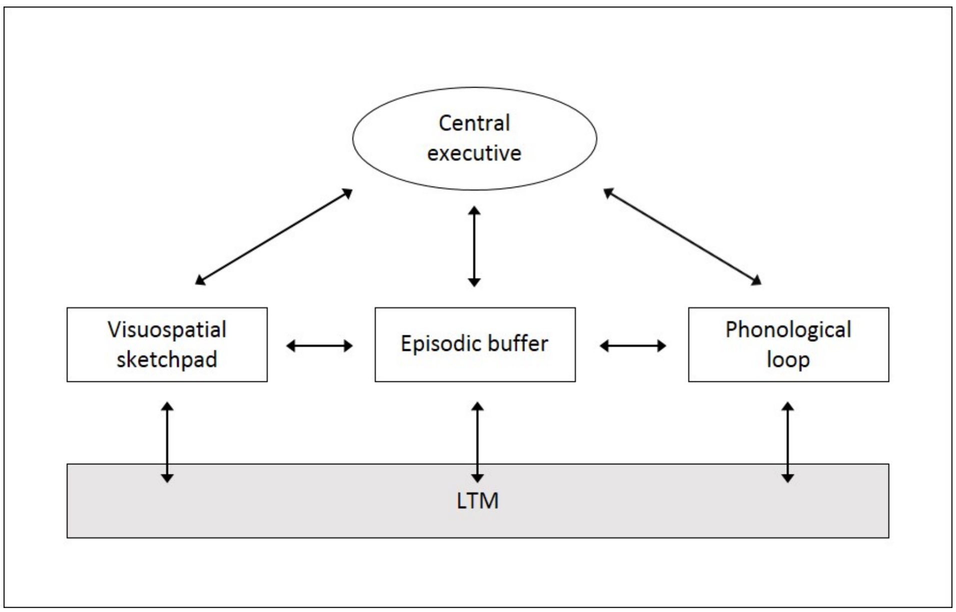


Figure 1- Four component model of working memory (Baddely 2007)

A feature common to all four components is that their capacity is limited in nature. This limitation is the key point in understanding the effect on user performance of a high load on working memory. In the following paragraph each component is briefly introduced.

The phonological loop: This system is responsible for storing phonological and speech-based information. This storage system is temporary and can be retained if it is refreshed by rehearsal. There is evidence that our brain stores phonological information in a separate memory system (Baddeley, 2007). For instance, Conrad & Hull (1964) found that the sequence of phonologically similar letters such as B D T G C P were remembered less precisely than a dissimilar sequence such as F K Y W R Q. Thus, other than storing purely acoustic information such as a tone, voice or music, our brain also stores the phonologic information of letters, words, and sentences in the phonological loop.

The visuospatial sketchpad: Visual and spatial information is stored in this system. It performs the same type of storage as the phonological loop but for visual information (Postle et al., 2006, Baddeley, 2007).

The central executive: This is the most important component of the working memory system as it performs the vital processing of coordinating and linking sub-systems. (Baddeley, 2007) identified four executive functions which are performed by the central executive component, "the capacity to focus attention, to divide attention, to switch attention and to provide a link between working memory and long-term memory". In this sense it is similar to the supervisory attentional system (SAS) model proposed by Shallice (1988). The central executive component possesses the capacity to direct and focus attention, which is the most fundamental function of working memory. Research on SAS and the central executive have strongly supported the fact that this capacity is limited, thus restricting the processing ability of the brain in any context.

Generally, our behavior is controlled by two mechanisms. The first is habitual action which relies mostly on long term memory (Norman & Shallice, 1986). The second mechanism activates when our brain perceives a deviation from a habitual process and needs to conduct a novel behavior Norman and Shallice (de Guinea & Webster, 2013). The latter mechanism is heavily dependent on central executive capacity. In situations where habitual responses are not relevant, the central executive is responsible for finding new solutions. Therefore, our performance is limited by the capacity of the central executive.

Three factors affect loading on the central executive. 1- Task complexity: complex tasks demand more working memory resources; performance is impaired as the load on working memory components increases. (Baddeley, 2007). 2- Degree of practice: as people gain more experience in performing a task, the demand on their working memory decreases (Baddeley, 2007). This happens because the mechanism for coordinating attention, thought, and behavior shifts gradually to one of habitual action which employs long term memory. In an information systems context, the degree of a user's experience with the the IT artifact and with the task itself, influences the level of cognitive load. 3-Task switching: research on this topic shows that there is an attentional cost when people switch between tasks, and it can heavily weaken their performance (Baddeley, 2007).

The episodic buffer: This system acts as an interface between all subsystems within working memory and long-term memory. Information from sub-systems and long-term memory will be integrated in the episodic buffer for processing by the central executive. The episodic buffer is accessible via conscious awareness. Thus, working memory needs the episodic buffer in order to have any access to long-term memory. As stated before, working memory is more than just a temporary storage; it is an interface that controls behavior and performs temporary manipulation and storage. It connects perception and memory, and attention and action (Baddeley, 2007). Given this crucial role, it affects perception, behavior, memory, and emotion.

Within working memory, there exist mental processes beyond users' conscious awareness (de Guinea et al., 2014). Although a portion of functions carried out by sub-systems of working memory is controlled through conscious awareness, there are also implicit processes ongoing in working memory (Baddeley, 2007). These processes are difficult to capture using a self-perceived measure since by definition they are beyond the awareness of people. Although these processes are implicit, they affect behavior, perception and emotion (Baddeley, 2007). For instance, subjects' perception of words will be affected by the brief presentation of a prior word. Subjects denied hearing the word money before presenting the ambiguous word bank, although they perceived it as a financial institution rather than the edge of a river. The effect of implicit processes has been strongly supported in relation to both working memory and long term-memory for a wide range of stimuli. Thus, in studying the effect of working memory on behavior, emotion and perception, both explicit and implicit processes must be taken into account. Finally, neuroimaging studies have well supported the fact that working memory processes are located in the frontal lobe (Roberts, Hager, & Heron, 1994).

Given the above, this thesis aims at introducing new measurements of cognitive load and the use of them as criteria for evaluating IT artifacts. We designed three experiments to examine the measurement of the cognitive load construct, and then to use them to evaluate the design elements of an online shopping website based on various types of cognitive load measures.

In the first essay, we designed two experiments to address both the challenge of measuring cognitive load and also evaluating an IT artifact against the newly developed measure. In the first experiment, we measured the user's instantaneous cognitive load during an online shopping task and extracted three features from it: Average load, Accumulated load, and Peak load. The traditional subjective cognitive load was measured as well in order to compare it with other metrics. Two task factors (task difficulty and task uncertainty) were manipulated in an online shopping task to test their effect on four different types of cognitive load (three implicit and one explicit). The results show that while all four measures were sensitive to task difficulty, accumulated load was the only measure that could capture the effect of task uncertainty. We therefore conclude that accumulated load is the most comprehensive measure among these four because it simultaneously captures both cognitive load and time dimensions.

The second experiment of the first essay uses accumulated load to evaluate the convenience of an IT artifact. We studied the effect of search convenience on the user's accumulated load and the effect of the latter on user satisfaction with the shopping experience. The literature suggests that inconvenience has two consequences: high workload and longer time on task (Jiang et al., 2013). Both of these factors can be captured by the accumulated load construct (Xie & Salvendy, 2000), so we hypothesize that convenience is negatively associated with accumulated load. As cognitive load increases, less working memory resources will remain for other mental processes, resulting in a less satisfying shopping experience. Thus, we hypothesize a negative link between accumulated load and satisfaction with the shopping experience. A single-factor (search convenience) experiment was designed to test the hypotheses, and three levels of search convenience were manipulated using three search functions. The results show that convenience is negatively linked to accumulated load, confirming our expectation that low convenience results in higher accumulation of cognitive load over time. Our analyses also show that accumulated load negatively influences users satisfaction with the shopping experience. It links accumulated load to an identified business need of users in the online shopping context.

Event-Related-Potential (ERP) is a technique developed to increase the reliability of EEG (Luck, 2012) by presenting an event several times to users and measuring their reaction to the stimulus in terms of the user's EEG. In the second essay, we design an ERP experiment in a natural setting where users have to perform a shopping task. More specifically, we investigate the contingent effect of product-sorting and the user's goals on user cognitive load. Products can be sorted based on different attributes (price, value, brand name, etc.). We argue that different types of sorting will result in cognitive load reduction if the user's criteria for choosing a product matches that attribute. For instance, users who are looking for a cheap product will have an easier shopping experience if the products are sorted based on price. We also hypothesize the effect of cognitive load on shopping performance. The less the cognitive load the more cognitive resources are available to users to process the task at hand, which means better performance. To measure cognitive load two different analyses were performed: ERP and frequency analysis. The ERPs were generated at two distinct events: 1- the moment that the task was presented to the user and 2- the moment that the user clicked on the target product. The frequency analysis was also performed to provide multiple evidence for the link between the match variable (i.e. between product sorting and user's goal) and cognitive load. The results strongly support the link between the fit construct and multiple cognitive load measures. Our results also support the link between cognitive load and task performance.

This thesis is an instance of NeuroIS potential to inform IS research. It is expected to contribute to research by advancing methods of measuring the cognitive load construct. The relevance of cognitive load for IS research had been emphasized in the past, however, the difficulty of measuring this construct stalled researchers in practically testing cognitive load theories. Using brain imaging tools, the two essays of this thesis provide various methods of measuring the cognitive load construct in an authentic human-computer interaction context. We discuss in detail how these measures differ from each other and how they can be used in IS research.

The first essay of this thesis provides an explanation for the causal link between convenience and one of the cognitive load measures (i.e. accumulated load). Convenience had been studied in the online shopping literature (Jiang et al., 2013), but never examined as an antecedent of consumers' accumulated load. The unique feature of the accumulated load construct allows us to link it to consumer convenience and also study it as a predictor of user satisfaction with online shopping.

In the first essay, we establish the accumulated load construct as a criterion for evaluating the convenience of shopping websites. The evaluation process of design science research requires researchers to test an existing IT artifact design against an identified business need (Hevner, 2007). By linking accumulated load to convenience and user satisfaction, this essay introduces accumulated load as one of the business needs that can be used to evaluate IT artifact design.

This thesis contributes to methodology by advancing the current experimental design practices in NeuroIS and IS. In the second essay, an Event-Related-Potential experiment is designed, within which one of the user's natural activities is used as an event to generate ERP waves. This practice can be replicated for other phenomena of interest in a user-IT interaction context.

This thesis contributes to practice by providing user-experience practitioners with a powerful tool for evaluating user interfaces. The first essay provides designers with a method to evaluate how a user's cognitive load varies in an authentic user-IT interaction context. This technique is useful not only because it precisely measures the user's' cognitive load, but also because it has high temporal resolution which enables designers to understand what users experience when they encounter every detail of the IT design.

The second essay shows that IT artifacts need to provide users with the right sequence of information, otherwise their cognitive load is increased and performance decreased, possibly resulting in their satisfaction being reduced when using the IT artifact. More specifically, online consumers have to be given the right sorting feature according to their shopping preference.

References

Baddeley, A. (2007). Working memory, thought, and action (Vol. 45). OUP Oxford.

BROCKE, J. V., Riedl, R., & LÉGER, P. (2013). APPLICATION STRATEGIES FOR NEUROSCIENCE IN INFORMATION SYSTEMS DESIGN SCIENCE RESEARCH. Journal of Computer Information Systems, 53(3).

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. British Journal of Psychology, 55(4), 429–432.

de Guinea, A. O., Titah, R., & Léger, P.-M. (2014). Explicit and Implicit Antecedents of Users' Behavioral Beliefs in Information Systems: A Neuropsychological Investigation. Journal of Management Information Systems, 30(4), 179–210.

de Guinea, A. O., & Webster, J. (2013). An Investigation of Information Systems Use Patterns: Technological Events as Triggers, the Effect of Time, and Consequences for Performance. Mis Quarterly, 37(4), 1165–1188.

DeStefano, D., & LeFevre, J.-A. (2007). Cognitive load in hypertext reading: A review. Computers in Human Behavior, 23(3), 1616–1641.

Dimoka, A., Pavlou, P. A., & Davis, F. D. (2011). Research commentary-NeuroIS: the potential of cognitive neuroscience for information systems research. Information Systems Research, 22(4), 687–702.

Gwizdka, J. (2010). Distribution of cognitive load in web search. Journal of the American Society for Information Science and Technology, 61(11), 2167–2187.

Hevner, A. R. (2007). A three cycle view of design science research. Scandinavian Journal of Information Systems, 19(2), 4.

Jiang, L., Yang, Z., & Jun, M. (2013). Measuring consumer perceptions of online shopping convenience. Journal of Service Management, 24(2), 191–214.

lo Storto, C. (2013). Evaluating ecommerce websites cognitive efficiency: An integrative framework based on data envelopment analysis. Applied Ergonomics, 44(6), 1004–1014.

Luck, S. J. (2012). Event-related potentials.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. Consciousness and self-regulation: advances in research and theory, ed. by RJ Davidson, GE Schwarts and D. Shapiro, 1–18. New York, NY: Plenum Press.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. Educational Psychology Review, 24(1), 27–45.

Postle, B. R., Idzikowski, C., Sala, S. D., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. The Quarterly Journal of Experimental Psychology, 59(1), 100–120.

Riedl, R., Davis, F. D., & Hevner, A. R. (2014). Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. Journal of the Association for Information Systems, 15(10), I.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. Journal of Experimental Psychology: General, 123(4), 374.

Schmutz, P., Roth, S. P., Seckler, M., & Opwis, K. (2010). Designing product listing pages—Effects on sales and users' cognitive workload. International Journal of Human-Computer Studies, 68(7), 423–431.

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge University Press.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. Cerebral Cortex, 6(1), 11–20.

Todd, P., & Benbasat, I. (1992). The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. Mis Quarterly, 373–393.

Todd, P., & Benbasat, I. (1999). Evaluating the impact of DSS, cognitive effort, and incentives on strategy selection. Information Systems Research, 10(4), 356–374.

Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science, 3(2), 159–177.

Xie, B., & Salvendy, G. (2000). Prediction of mental workload in single and multiple tasks environments. International Journal of Cognitive Ergonomics, 4(3), 213–242.

Chapter 2-
Essay I - Accumulated Cognitive Load as a Criterion for Evaluating IT Artifacts

Abstract

Two studies were conducted to investigate how users' accumulated cognitive load can be used to evaluate the convenience of an IT artifact, defined as any interface element that adds to users' comfort. The first study addressed the challenge of measuring cognitive load. An experiment was designed to manipulate two antecedents of convenience (task difficulty and task uncertainty) in an online shopping context. Users' instantaneous cognitive load was measured using electroencephalography (EEG). Three features of cognitive load, including accumulated load (the sum of instantaneous cognitive load over time) were then extracted from the EEG data and compared against self-reported cognitive load. Results suggest that, although all four cognitive load measures (accumulated load, peak load, average load, and self-reported load) are sensitive to task difficulty, only accumulated load is able to capture task uncertainty. In the second study, we investigated the effect of search convenience on users' accumulated load and the effect of the latter on user satisfaction in an online shopping context. A between-subject experiment was designed. Results suggest that convenience negatively influences accumulated load, and the latter negatively influences user satisfaction. Overall, our findings show that accumulated cognitive load can be used to accurately evaluate the convenience of IT artifacts.

1.1 Introduction

Design science research deals with the design and evaluation of IT artifacts. In the *evaluation process* of design science research, an IT artifact (e.g., online grocery shopping website) is tested against an identified business need (e.g., user cognitive load) (Hevner, 2007). Brocke et al. (Brocke, Riedl, & Léger, 2013) proposed three strategies for the use of neuroscience in design science research. The first strategy deals with the use of neuroscience theories to inform the building and evaluation of IT artifacts. The second strategy includes using neuroscience tools to test IT artifacts, and finally the third strategy is the use of neuroscience tools as built-in elements of IT artifacts. Neuroscientists have progressed in explaining the neurophysiological basis of a number of constructs that are interesting for us in the IS field, such as trust, cognitive load, and emotions (vom Brocke & Liang, 2014). Therefore, this knowledge base can be used to study design problems and find solutions to them. For instance, electroencephalography (EEG) can be used to monitor users' level of cognitive load and emotion during their interaction with an IT artifact (Brocke et al., 2013). Although users' cognitive load can be theoretically linked to design criteria, to the best of our knowledge it is not yet considered as a criterion for designing IT artifacts in research and practice.

Cognitive load is considered as the cost of decision making (Todd & Benbasat, 1999) and reflects how efficiently a task is performed by a human brain. Users prefer to minimize this cost in order to have a more satisfying experience with an IT artifact (lo Storto, 2013). Cognitive load has been found to be an antecedent of a wide range of behavioral and perceptual constructs (Baddeley, 2007), such as user performance, negative emotions, fatigue, and satisfaction (Garbarino & Edell, 1997; Gwizdka, 2010). For instance, users' satisfaction with a shopping task is negatively associated with the amount of cognitive load they spend on the task (lo Storto, 2013). Therefore, it is important to find ways to evaluate IT artifacts based on users' cognitive load. To achieve this goal, two challenges need to be overcome. First, one needs to measure cognitive load precisely. Cognitive load is a temporal variable, meaning that users have different levels of cognitive load depending on the moment. Thus, cognitive load measurement is about not only validity and reliability of the measures, but how we aggregate these measures over time. A highly temporal measurement of cognitive load provides users' instantaneous level of cognitive load at any moment during a user-IT interaction. However, running typical statistical analysis requires to aggregate these measures into a single value. The type of aggregation (e.g., maximum, minimum, average, sum) may result in very different outcomes.

Second, an appropriate theoretical lens needs to be chosen in order to explain how cognitive load can improve IT artifact design. The theoretical framework should be able to link well-developed design concepts to users' cognitive load. We designed two experiments to address the measurement challenge and also propose a theoretical model to explain how cognitive load relates to design related concepts.

Measuring a cognitive load construct has been a challenging task for researchers across different fields (Gopher & Braune, 1984; Paas & Sweller, 2012; Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Cognitive load is among the complex cognitive processes that people are unable to self-report truthfully (vom Brocke & Liang, 2014). It encompasses different aspects which are beyond users' consciousness, meaning that using the traditional self-reported scales are not sufficient to capture all aspects of this construct (de Guinea, Titah, & Léger, 2014). Developments in the field of neuroscience have provided researchers in other fields such as marketing and information systems with new tools and methods to capture a wide range of cognitive, emotional, and behavioral constructs.

The use of neurophysiological measures of cognitive load can alleviate these shortcomings to some extent. Researchers have compared self-reported and neurophysiological measures of mental workload and suggested that the latter can provide a more comprehensive and richer understanding of the workload construct (De Guinea, Titah, & Léger, 2013). However, thus far, research has only used some of many possible measures of implicit workload when comparing it to self-reported workload (Xie & Salvendy, 2000). Given that multiple measures of neurophysiological workload can be extracted from EEG signals (e.g., Average load, Accumulated load, and Number of peaks), an even richer understanding of cognitive load could be attained.

In our first study, we use EEG to measure users' instantaneous workload, i.e. users' cognitive load at any time, and extracting various types of cognitive load measures from it. Two task characteristics (i.e., difficulty and uncertainty) are manipulated in an online shopping environment to capture different types of cognitive load. Task difficulty and uncertainty are two factors that increase the time and effort consumers need to perform a shopping task, which consequently reduce their shopping convenience (Jiang, Yang, & Jun, 2013). We find that one of the measures extracted from instantaneous cognitive load, called accumulated load can be used as a proxy to evaluate the convenience of an online shopping experience.

The second study is designed to use accumulated load to evaluate the convenience of a shopping website. Users' convenience refers to any factor during the online shopping process that adds to their comfort (Moeller, Fassnacht, & Ettinger, 2009). We argue that accumulated load reflects users' comfort in the shopping process because it measures two resources that are associated with users' convenience: mental effort and time. Our results show that accumulated load is influenced by user convenience, and predicts user satisfaction.

This research makes several contributions to theory, and also has implications for practice. First, it addresses the challenge of measuring cognitive load, proposing a valid EEG-based cognitive load measure that captures relevant IT interaction constructs such as task difficulty and user satisfaction. We use EEG to measure users' instantaneous cognitive load, which can be used to generate other types of cognitive load measures. These measures, although extracted from the same baseline measure (i.e., instantaneous cognitive load), represent different aspects of the cognitive load construct. Average load is the mean of users' instantaneous load over time, and gives an overall measure of a user-IT interaction. Peak load shows the highest cognitive load experienced by users during a transaction, and can be used to understand the inefficient elements of an IT artifact. This contribution is in line with the first strategy proposed by Brocke et al. (Brocke et al., 2013) in the use of neuroscience to inform the building and evolution of IT artifacts.

Second, our paper contributes to theory by proposing a causal link between consumer convenience and accumulated load. This contribution also matches the first strategy suggested by Brocke et al. (Brocke et al., 2013). This strategy requires that mental processes and perceptions be linked to human brain activities. Our study explains how the accumulated load is associated with the consumers perception of convenience.

Third, and in line with the second strategy proposed by Brocke et al. (Brocke et al., 2013), we evaluate an IT artifact based on users' accumulated load. This strategy includes the use of neuroscience measures to evaluate IT artifacts. Our research evaluated the convenience of an existing IT artifact against users' accumulated load. The results, suggest that accumulated load can be used as a tool to evaluate IT artifacts.

Our research also has implications for practice by providing a metric for designers for assessing the convenience of an IT artifact and increasing user satisfaction. Accumulated load could be used by designers to identify the inconvenient parts of a user-IT transaction and evaluate the modified IT artifact against the old ones.

1.2 Literature Review

Cognitive Load

Cognitive load can be defined as the set of working memory resources used to perform a task (DeStefano & LeFevre, 2007). Despite the fact that there are different definitions of cognitive load in the literature (Baddeley, 2007; Todd & Benbasat, 1999; Wickens, 2002), there are two common elements among them which are the essential components of the cognitive load construct: 1- working memory resources and 2- the interplay between mental resources and task demands. Theories developed in the fields of cognitive psychology and neuroscience state that working memory is responsible for a wide range of processing tasks and short term storages (Baddeley, 2007; Smith, Jonides, & Koeppe, 1996). Working memory has four components, which include three storage systems and a central executive unit. The three storage systems (i.e., visuospatial sketchpad, episodic buffer, and phonological loop) are responsible for temporarily retaining different types of sensory information. The central executive unit, which is the most important component, performs the crucial task of coordinating and linking working memory subsystems and attention. Therefore, task demands that require users to encode, activate, store, and manipulate information are imposing cognitive load on their working memory (DeStefano & LeFevre, 2007).

All components of working memory have limited resources which bound the processing and storage capabilities of the human brain (Baddeley, 2007; Liang, Peng, Xue, Guo, & Wang, 2015; Wickens, 2002). Researchers from various disciplines such as education, psychology, and management studied how cognitive load affects different types of individual performance (De Jong, 2010; Ryu & Myung, 2005; Wickens, 2002). For instance, in education, students' learning performance is found to be affected by their level of cognitive load (Paas, Renkl, & Sweller, 2003). Cognitive load theory has conceptualized three types of cognitive load with respect to students' learning performance. Intrinsic load, which refers to the load generated by the interaction between the nature of the instructional material and the learners' expertise (Sweller, 1994). Extraneous load is the ineffective load imposed by ill-designed instructional materials. Finally, germane load is the amount of load that contributes to the main mechanisms of learning (Paas, Renkl, et al., 2003). Cognitive load theory suggests that in high cognitive load situations, extraneous load should be decreased to enhance students' learning performance (Sweller, 1994).

The link between performance and cognitive load has been studied in the information systems field as well. Users perform well as long as their level of cognitive load is within their cognitive capacity limit (Liang et al., 2015). Exceeding this level results in cognitive overload, meaning that there are not enough cognitive resources to process the information required to perform the task. For the same reason, users cannot pay attention to details when experiencing high cognitive overload, thus, they make poor decisions even though the information is clearly presented (Minas, Potter, Dennis, Bartelt, & Bae, 2014).

Cognitive load has been linked to efficiency in decision making; it is conceptualized as the cost of decision making (Hoque & Lohse, 1999). This cost, which refers to the amount of cognitive resources used by the brain to make a decision, should be minimized to achieve efficiency in decision making. Reducing this cost is one of the two ways that decision aid tools help users (the other is increasing decision quality) (Todd & Benbasat, 1992).

The limitation of working memory resources not only affects user performance and efficiency but users' perception and cognition (Baddeley, 2007). Research shows that high cognitive load results in experiencing negative consequences such as frustration, negative affect, or mental fatigue (Mizuno et al., 2011). Lower cognitive load has also been related to users' satisfaction with online tasks(Gwizdka, 2010). In the online shopping context, consumers are more satisfied with websites that show only the necessary information and avoid presenting excessive information to users (lo Storto, 2013). It is thus important to identify and study the elements, for instance in an online shopping task, that contribute to users' cognitive load.

Cognitive overload can be a result of either a task or system demand or both. For instance, cognitive overload due to problematic use of social networking sites worsens the academic performance of students (Turel & Qahri-Saremi, 2016). System elements can, on the one hand, be the source of cognitive overload or, on the other hand, help users to decrease their level of cognitive load leading them to perform better on a task. For instance, online consumers sometimes have difficulty in selecting a product when there are many alternatives; they need to gather, screen, and evaluate product information, and this can be cumbersome if the consideration set is relatively large (Sénécal, Léger, Riedl, & Davis, 2018). Product recommendation agents can assist users in processing information about products , and spare them with cognitive overload (Aljukhadar, Senecal, & Daoust, 2012; Qiu & Benbasat, 2009). Aljukhadar et al. (Aljukhadar et al., 2012) suggest that information overload influences consumers' decision making strategies. As information load increases on a website, consumers rely more on recommendation agents to cope with the situation. Group support systems play a similar role in facilitating collaboration among members thus decreasing their cognitive load level (Briggs, De Vreede, & Nunamaker Jr, 2003).

Cognitive Load Measurement

There are three types of cognitive load measures: self-report, performance based, and neurophysiological (Gopher & Donchin, 1986; O'Donnell & Eggemeier, 1986). Self-report measures, which are the most mature ones, have been used for many years to study users' behavior (Colle & Reid, 1999). In this type of measure, users are asked to reflect on their experience of the task or system, and report the level of difficulty they experienced in their interaction with the task or system.

With performance-based measures, users' performance on the task they perform is used as an indicator of their workload level. There are two variations of this type of measurements. The performance could be measured on either a) the primary task or b) a secondary task that interrupts the primary task. This measurement approach assumes that users' performance on a task is a direct measure of task difficulty, which implies the amount of workload used to perform that task (Gopher & Donchin, 1986).

Finally, signals generated by different parts of the body can be used to draw inference about users' cognition and emotion. Psychophysiological tools such as eye tracking, skin conductance, and brain imaging tools such as EEG and fMRI lie in this category of measures (Dimoka et al., 2012). A summary of different types of measures can be found in Table 1. For a review of methodological issues relating to different measurement tools see (Dimoka et al., 2012; Gopher & Donchin, 1986).

Table 1- Cognitive Load Measurement

Measurement Method	Description	Example	Strength	Weakness	Example from IS Research
Self-reported	Users are asked to report the experience they had with the task and the system	Questionnaires	Face validity	-Consciousness assumption - Retrospective bias	Hender et al. (2014) Blohm et al. (2016)
Performance	Measuring users' workload based on their performance on the primary task or a secondary task	Using participants' performance on a secondary arithmetic task as a measure of their workload on a primary task (e.g. online shopping)		-Assumption of association between performance and workload	
Neurophysiological	Measuring users' workload based on various neurophysiological metrics	-Pupil dilation -Cardiovascular measures -Electroencephalography	-Measuring beyond users' consciousness -Temporal resolution	-Artificial Setting - Content Validity	Minas et al. (2014) Ortiz de Guinea et al. (2013)

Borrowing from neuroscience and neuropsychology, researchers in the Management field have started to use neurophysiological instruments to measure cognitive load (Gwizdka, 2010). Researchers have developed several algorithms based on EEG signals to calculate mental workload metrics (Coyne, Baldwin, Cole, Sibley, & Roberts, 2009). EEG signals are high dimensional noisy time series (e.g., 500 data points per second), which encompass a high volume of information (Garrett, Peterson, Anderson, & Thaut, 2003). In order to relate EEG signals to specific mental states (e.g., mental workload), first of all, the signal has to be cleaned (i.e., noise and artifacts should be removed), and more importantly, relevant signal features that represent the desired mental state should be extracted (Brouwer, Zander, van Erp, Korteling, & Bronkhorst, 2015). In this research, we define four types of cognitive load, which reflect different aspects of human cognitive load (Figure 1).

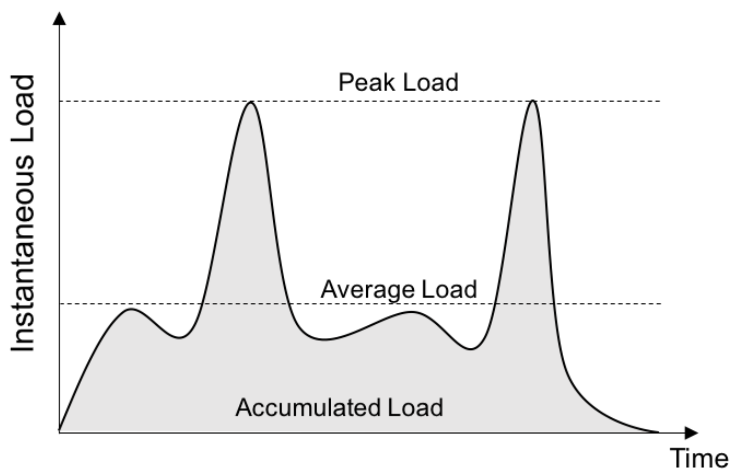


Figure 2- Workload Features

Xie and Salvendy (Xie & Salvendy, 2000) define four types of workload based on users' instantaneous load: peak load, average load, accumulated load, and self-reported load. Instantaneous load shows users' cognitive load at any given moment (see Figure 1). Peak load is the maximum load an individual experiences during a task, while accumulated load is the summation of all the cognitive load. Average load reflects the mean load an individual experiences in performing a task. Self-reported load is the amount of cognitive load reported by users after performing a task (Xie & Salvendy, 2000). Figure 1 illustrates how peak load, accumulated load, and average load are related to instantaneous load. Despite the fact that all workload measures represent users' cognitive load level, they are nevertheless different. Average load is a general measure which represents the amount of cognitive resources that a task requires on average. Accumulated load shows the whole amount of load that is experienced taking into account the time needed to perform a task; it includes a time dimension in addition to instantaneous load measures. Peak load yields unique information about users' cognitive load when interacting with a task/system; it cannot be obtained using a self-reported measure, average load, or accumulated load. We know from the literature that cognitive capacity is limited (Wickens, 2002) and that there will be negative consequences (e.g., decline in performance and satisfaction) if users exceed this limit (Colle & Reid, 2005). Apart from accumulated load, workload measures do not reflect the temporal aspects of cognitive load. This information can be useful for studying users' coping behavior with high workload situations. For instance, if users experience high cognitive load during a difficult task they might withdraw from the task (Venables & Fairclough, 2009).

Shopping Convenience

Convenience can be defined as any element in the shopping process that adds to a consumer's comfort (both physical and mental) (Jiang et al., 2013). There are two factors associated with a convenient shopping activity: 1) time saving and 2) effort minimization (Seiders, Berry, & Gresham, 2000). Berry et al. (Berry, Seiders, & Grewal, 2002) suggest that time costs are negatively associated with consumers' perception of convenience. For example, the longer the online shopping process, the less consumers' convenience would be. Berry et al. (Berry et al., 2002) also point out that convenience perception is negatively linked to consumers' perception of cognitive, physical, and emotional effort. Therefore, any factor (e.g., design element) that induces effort in users, will decrease their convenience, and consequently their satisfaction (lo Storto, 2013).

Convenience is analogous to perceived ease of use as both concepts reflect the effort needed to perform a task using an information system; however, for the purpose of our study, we suggest that the convenience construct is more relevant for five main reasons.

1. Although convenience and PEOU are conceptually similar, research findings show that convenience is an antecedent of PEOU (Kim, Mirusmonov, & Lee, 2010). Convenience explains the amount of time and effort required by an IT artifact to perform a task, thus more convenience results in perceiving more ease of use.
2. The convenience construct developed by Jiang et al. (Jiang et al., 2013) is a well-suited lens for studying design elements because it is developed based on shopping sub-processes that reflect website elements. Jiang et al. (Jiang et al., 2013) break down consumer convenience in online shopping environments into several dimensions: a) access convenience, b) search convenience, c) evaluation convenience, d) transaction convenience, and e) possession/post-purchase convenience.

An explanation of each convenience dimension is provided in Table 2

Table 2- Convenience dimensions based on Jiang et al. (2013)

Convenience dimension	Definition	Example
Access convenience	It deals with the accessibility of the shopping website	- Availability of the website at any time from any location - Availability of products and brands
Search convenience	It concerns any factor that facilitates or hinders finding a product.	- Website navigation - Search function - Product classification
Evaluation convenience	It is related to the availability of detailed product description so that consumers can compare products.	- Product information - Product categorization
Transaction convenience	It concerns easy payment methods and finishing a shopping transaction on a website.	- Check out process - Payment methods
Possession/post-purchase convenience	It is about the convenience of receiving the product and any possible post-purchase issue.	- Delivery - Product undamaged

3. Consumer convenience has been studied in marketing and found to be a predictor of several dependent variables such as customer satisfaction (Colwell, Aung, Kanetkar, & Holden, 2008), behavioral intentions (Seiders, Voss, Godfrey, & Grewal, 2007), online service quality, customer service, and trust (Colwell et al., 2008). Research suggests that convenience is a major factor in intensifying consumers' relationships with a service provider and inconvenience is found to be one of the reasons that consumers exit such relationships (Moeller et al., 2009).
4. It is closely related to users' level of cognitive load. Berry et al. (Berry et al., 2002) suggest that consumers' perception of convenience is negatively associated with their perception of cognitive and emotional effort. Therefore, the conceptual

relationship between convenience and cognitive load has already been acknowledged in the literature.

5. Convenience captures a wide range of costs associated with online shopping. Any factor that contributes to the extent to which consumers avoid time and effort in an online shopping task influence their convenience (Moeller et al., 2009). For instance, information uncertainty reduces the quality of information provided to consumers, which negatively influence the evaluation convenience of the shopping process (Jiang et al., 2013). Task difficulty (e.g., difficulty to assess the quantity to purchase) increases the time and mental effort required to perform the shopping task and reduces the shopping convenience.

In Study 2, we examine how search convenience affects users' cognitive load and satisfaction. Thus, in the next section we briefly define the dimensions of online search convenience.

Search Convenience

Jiang et al. (Jiang et al., 2013) suggest that consumers mention search inconvenience and difficulty in finding the desired product is one of the major obstacles to efficient online shopping. They found that search convenience accounts for the largest portion of explained variance (31%) among the five dimensions of online shopping convenience. Any barrier related to searching and finding a product in the website is categorized under the search convenience construct. Search inconvenience factors can be grouped into four major categories: 1) Download speed, 2) Website design 3) Search functions, and 4) Product classification (Jiang et al., 2013).

Download speed relates to the quality of the internet connection. Website design represents the extent to which it is difficult to navigate through website pages and understand its structure. Search functions explains how fast and easy users can search for products and find what they want. Finally, Product classification relates to the right use of product categories and sorting in a way that makes finding a product easy. For instance, users on an online grocery website should be able to easily find the entry field to search for products or locate the website filters to search for organic products (Website design), search for multiple products and compare them (Search function), and use product categories to browse and reach their target products (Product categories)

Based on the literature of convenience and the abovementioned findings, we suggest that accumulated load is conceptually linked to the convenience concept. Accumulated workload has two primary dimensions: 1- overall level of workload experienced during a task and 2- the total time spent on a task (Paas, Tuovinen, et al., 2003). To calculate accumulated load, we need to sum users' instantaneous mental workload over time. It thus takes into account both *mental effort* and *time costs*, which are two primary factors of consumers' convenience (Berry et al., 2002). We argue that convenience influences accumulated load. This is not the case with the other measures of cognitive load (Average load, peak load, and self-reported load) because they do not include the time dimension, and consequently represent different aspect of the cognitive load construct. Our two studies are designed to examine the effect of user convenience on accumulated load and user satisfaction. The first study addresses the measurement challenge of cognitive load. We measure users' instantaneous workload by manipulating two task factors (difficulty and uncertainty) and extract three features (average, accumulated, and peak) from it. Task uncertainty and difficulty are both determinants of online shipping convenience. In the second study, we investigate the relationships between convenience, accumulated load, and user satisfaction.

1.3 Study 1: Cognitive Load Measurement and Consumer Convenience

The objective of the first study is to measure instantaneous workload, extract three features from it (i.e. average load, peak load, and accumulated load), and study how two determinants of online consumer convenience, which induce workload, affect the three measures along with a self-reported measure of workload. A frequency analysis on users' EEG is performed to calculate their instantaneous cognitive load and then the three abovementioned features are extracted. We chose difficulty and uncertainty as the two task factors because they are both relevant in the context of online shopping and also determinants of online consumer convenience. The two factors are predictors of cognitive load, which enables us to address the challenge of measuring cognitive load.

Cognitive load imposed on users' working memory can come from different sources such as task characteristics or system demands. In general, any factor that makes decision making more difficult will impose more load on users' cognitive resources. For instance, research suggests that mathematical complexity and arithmetic operations increase users' mental workload (Ryu & Myung, 2005). Difficulty and cognitive load are so linked together that many studies ask participants to rate their perception of task difficulty as a measure of cognitive load (Gopher & Donchin, 1986). Task difficulty refers to the extent to which performing a task is difficult. It is the main driver of workload because task demands impose load directly on working memory resources. Therefore, we hypothesize that task difficulty will affect users' mental workload.

H1: Task difficulty is positively associated with all workload types (self-reported, average, accumulated, and peak loads).

Information processing includes three stages: information perception, decision/response selection, and response execution (Xie & Salvendy, 2000). The information flows from the first stage to the last where a decision is made and executed by the user. There are various factors that can impair this process and make the decision process more difficult for users. For instance, irrelevant information or information overload makes the first stage more difficult for the user to perceive the information. In a similar way, information uncertainty affects the first stage by making it more difficult for users to assess the critical information required for making product decisions, which consequently imposes more cognitive load on users (Aljukhadar et al., 2012).

Uncertainty can be classified into two types: 1- affective uncertainty, which relates to affective factors of the task/situation such as pessimism or optimism, and 2- cognitive state uncertainty, which relates to the ambiguities involved in rational decision making on the part of the user (Wilson, Ford, Foster, & Spink, 2000). Cognitive state uncertainty includes factors such as information ambiguity, meaning that the whole or part of the critical information required to make a decision is not presented clearly to the user. Under such conditions, more cognitive effort is required of the user to process the information and make a decision. Thus we hypothesize that task uncertainty increases users' cognitive load.

H2: Task uncertainty is positively associated with all workload types (self-reported, average, accumulated, and peak loads).

Method

Experimental Design and Procedure

To test our hypotheses, a 2 level task difficulty (Low (L) or High (H)) X 2 task level uncertainty (Low (L) or High (H)) within-subject experiment was designed. This experiment was approved by our Institutional Review Board (IRB). Ten subjects participated in the experiment and 50% were male. Each subject performed four tasks (HH, HL, LH, LL) which were randomly ordered. Participants first filled out a questionnaire and then moved on to the first shopping task. For each task, participants had to shop on a selected online grocery website. The task started on an online grocery recipe page. Participants were instructed to shop for five given items for each assigned recipe. After finishing each task, subjects completed a questionnaire assessing their level of cognitive load (self-perceived load). The experimental procedure is presented in Figure 3.

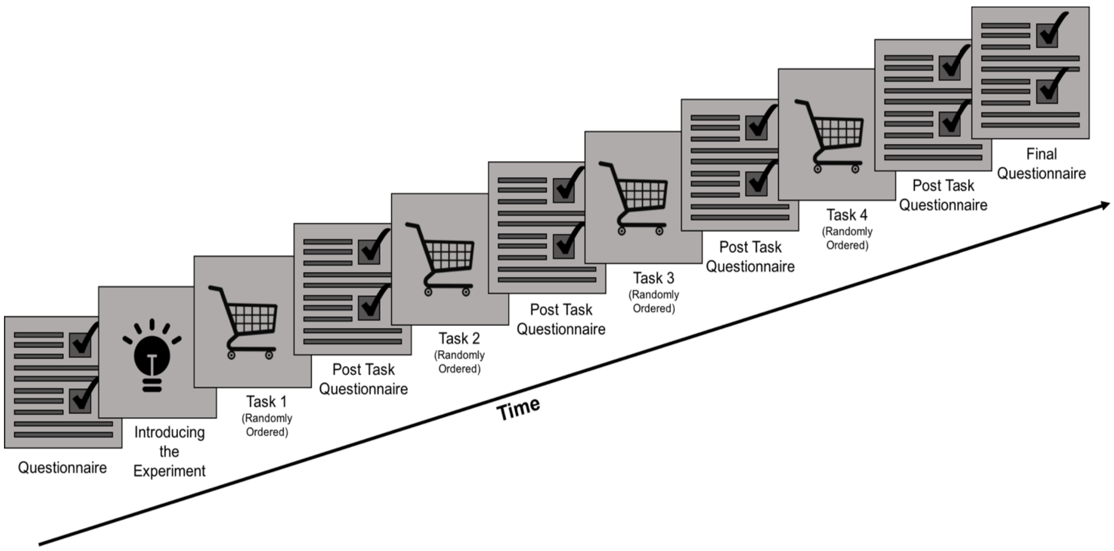


Figure 3- Experimental Procedure of Study 1

Online grocery is a suitable context for this study because it is a time-consuming user-IT transaction which includes several decision-making points (Desrocher, Leger, Sénécal, Pagé, & Mirhoseini, 2015). As opposed to other online shopping contexts which are mostly composed of single product decisions, online grocery usually includes several decisions and allows us to observe the dynamics of users' cognitive load. Moreover, there are a number of workload antecedents which are unique to this context such as performing simple arithmetic operations in order to find the right quantity of a product or the uncertainty involved in buying perishable products.

Task difficulty was manipulated by asking subjects to change the quantity of ingredients suggested for the recipe. Recipe pages suggested the right amount of each grocery item to use for 4 people. In the low task difficulty condition, participants had to choose the same quantity and in the high difficulty condition, they were asked to perform simple arithmetic operations in order to find the right quantity of each ingredient for 20 people.

We used product type (perishable or non-perishable) to manipulate task uncertainty. Based on transaction cost theory, Chintagunta et al. (Chintagunta, Chu, & Cebollada, 2012) propose nine types of costs involved in a decision to buy groceries. One of these costs is quality inspection or product evaluation cost. In online grocery shopping, consumers use product descriptions and product images to infer the quality of products. Comparing in-store and online grocery shopping, product image of non-perishable products (e.g., shampoo, box of chocolate) gives the same visual information about the product. However, for perishable products (e.g., meat, vegetables), consumers need to touch, smell, see, or even taste them before making a decision. Industry reports also show that the adoption rate of perishable products is the lowest among all product types in the online grocery sector[1]. Therefore, there is more uncertainty involved in product quality inspection of perishable products compared to non-perishable products. In the current experiment, participants had to shop for 5 specified non-perishable products (e.g., olive oil) in the low *uncertainty* condition, and 5 specified perishable products (e.g., meat) in the high *uncertainty* condition.

Measures

Self-reported cognitive load was measured using a five item scale developed by Cameron (Cameron, 2007). Instantaneous measure of cognitive load was measured using a linear EEG algorithm. In general, EEG oscillations are categorized into four frequency bands: Delta (0 to < 4 Hz), Theta (4 to < 8Hz), Alpha (8 to 13 Hz), and Beta (>13 Hz) (Libenson, 2012). In this experiment, we use a linear formula based on three frequency bands to measure the cognitive load of users. This algorithm includes calculating the ((delta+theta)/alpha) power ratio over a moving 2 second window and comparing it to the average of the previous 20 seconds (Coyne et al., 2009). The average of users' workload over a task period was used to represent *Average load*. The area under the instantaneous workload curve which equals the sum of instantaneous load over time was used to

[1] Nielsen Report, The Future of Grocery E-commerce, Digital Technology and Changing Shopping Preferences AroundThe World, April (2015).

calculate *Accumulated load*. To calculate the *Peak load*, we counted the number of times that the amplitude of instantaneous load exceeded 99% of the whole instantaneous load measures (2.5 standard deviations), indicating the number of times participants experienced nearly peak load. In other words, our measure of peak load counts the number of times the instantaneous load value lies in the top 1% of the total instantaneous load values.

Apparatus, Data Acquisition, and Analysis

EEG data was recorded using 32 electrodes using EGI's dense array electroencephalography (dEEG). To clean the data, three filters were applied in MATLAB to the EEG signals: a high pass filter at 1.5 Hz, a low pass at 50Hz, and a notch between 55Hz and 65Hz. Then the DC component of the signal was removed using *detrend* command in MATLAB, which is equivalent to removing the mean value from the signal vector. Artifact removal was performed in two steps. First a continuous EEG rejection function in EEGLAB toolbox was used. The function included a 500 ms moving window with steps of 250 ms that calculates the max-min voltage in each interval and removes the data if the voltage difference exceeds 400 uv. In the next step, independent component analysis (ICA) was performed on the signal in order to decompose the signal into 32 independent components. The ADJUST plugin in EEGLAB toolbox was used according to Mongnon et al. (Mognon, Jovicich, Bruzzone, & Buiatti, 2011) to identify artifacts such as eye blinks, eye movements, and generic discontinuities. After identifying and removing bad components, the signal was reconstructed in time domain for calculating instantaneous workload. We used Fast Fourier Transformation (FFT) to quantify the signal power based on four frequency ranges. Delta 0 to < 4 Hz, Theta 4 to < 8Hz, Alpha 8 to 13 Hz, and Beta >13 Hz. Then the ratio of (delta+theta)/alpha was calculated according to the procedure suggested by Mikulka et al. (Mikulka, Scerbo, & Freeman, 2002) and Charland et al. (Charland et al., 2015). The resulting curve was used as the instantaneous workload of each user.

In order to test the hypotheses, we used a regression analysis with task *difficulty* and *uncertainty* as independent variables and mental workload (Self-reported*, Average,* *Accumulated, Peak loads*) as the dependent variable. The data analysis was performed using the Stata software package (StataCorp, Texas, United States). Since the observations are non-independent, we used panel regression analysis to account for within-subject dependencies (Xtreg command in Stata). We also controlled for learning effect in the regression model by including the order in which participants performed the task.

Results

The manipulation checks show that both difficulty (M_{Low}=3.05, M_{High}=4.70, $p<.01$) and uncertainty (M_{Low}=2.85, M_{High}=3.40, $p<.05$) manipulations were perceived differently, as expected. H1 proposed that task difficulty is positively associated with all measures of mental workload. Results show significant relationships between difficulty and self-reported load (b= 0.88, $p<0.01$), average load (b=0.02, $p<0.05$), accumulated load (b=44.45, $p<0.01$), and the number of peaks (b=4.87, $p<0.05$). Therefore, H1 is supported. Our second hypothesis proposed that task uncertainty is an antecedent of all workload types. Regression results show that the relationships between uncertainty and self-reported load (b=0.51, p=0.18), average load (b=0.01, p=0.15), and number of peaks (b=3.12, p=0.15) are not significant. However, a significant result was found for the relationship between uncertainty and accumulated load (b=35.24, $p<0.05$). Thus, H2 is partially supported.

To better understand the difference between workload measures, we conducted a post hoc analysis and compared the effect of task *difficulty* and *uncertainty* on each workload measure. We used four regression models with the same independent variables (Task *difficulty* and *uncertainty*) but with four different measures of workload as dependent variable (Self-reported, accumulated, average, and number of peaks). Since the number of independent variables are the same, we compare R-square across the regression models. Accumulated load explained more variance than other models (21%). It was followed by number of peaks (16%), self-reported load (14%), and average load (5%).

Discussion

Our results suggest that all the extracted features of instantaneous load and the subjective measure of workload (i.e., self-reported load) are sensitive to task *difficulty*; however only accumulated load was able to capture the mental workload induced by task *uncertainty*. Based on the definition of accumulated load, we can expect more comprehensiveness of this measure compared to average load or self-reported load. Accumulated load measures the fluctuations in users' instantaneous workload over time, thus accounting not only for Self-reported load but also the total time that user has been performing the task. "Time on task" has been used as a measure for workload before (DeLeeuw & Mayer, 2008), thus it may capture more dimensions of the workload construct.

The post hoc analysis confirms our results, i.e., that accumulated load captures more variability caused by workload inductors. It also shows a relatively high R-square for number of peaks. In addition to being used as a workload measure, peak load can reflect the moments that users experience high workload. Research findings suggest that exceeding users' cognitive capacity limit results in a number of negative consequences (Liang et al., 2015). Peak load measurement enables us to identify the number of times that users approach their cognitive capacity limits. This type of information, which can only be gained using peak load, may provide a proper lens to study users' coping behavior in dealing with high workload situations.

This study shows that the total cognitive load experienced (i.e., accumulated load) in a task is sensitive to both task difficulty and uncertainty, showing that accumulated load is more powerful in capturing users cognitive load compared to the other three measures. Accumulated load dimensionality (i.e., time and instantaneous load) allows us to link it to other constructs in the online shopping literature that are conceptually related to time and mental effort. This leads us to the concept of online consumer convenience, which reflects users' comfort in online shopping including time costs and mental effort. In the next study, we examine how users' accumulated load is influenced by convenience.

1.4 Study 2: Accumulated Cognitive Load as a Criterion to Evaluate Information Systems Convenience

In line with the first and the second strategies proposed by Brocke et al. (Brocke et al., 2013), our second study aims at introducing the accumulated load construct as a criterion for evaluating IT artifacts. In this study, we use the measures tested in the first study to examine the search convenience of an IT artifact. Search convenience includes all activities that users perform in the process of finding a product. Any difficulty that results from search convenience factors (i.e. search function, website design, product categories, and download speed) is considered as search inconvenience. Inconvenience causes more cognitive load and more time spent on a task both of which are dimensions of accumulated load (Paas, Tuovinen, et al., 2003). Therefore, we expect that the less convenient the IT artifacts is, the more cognitive load users experience and the more time they need to perform a task using that artifact. Hence:

H3: Search convenience negatively influences users' accumulated cognitive load.

Website interfaces should provide users with the information necessary for shopping, and prevent them from being overloaded with excessive information (lo Storto, 2013). Ill-designed websites occupy the user's working memory resources more than necessary either by failing to provide critical information for decision making or by providing redundant information, leading to an unsatisfying shopping experience (lo Storto, 2013). In a user-IT transaction, high accumulated cognitive load means that the user has experienced a relatively high cognitive load for a long period of time, which leads to the user feeling frustration and negative affects (Mizuno et al., 2011). This may affect the emotional aspect of satisfaction (Spreng & Mackoy, 1996) and strengthen the negative relationship between accumulated load and satisfaction.

H4: Accumulated load negatively influences user satisfaction.

In the literature, it is suggested that convenience has a direct positive relationship with satisfaction (lo Storto, 2013). However, as mentioned before, search inconvenience has two consequences, increasing cognitive load and shopping time, both of which are dimensions of the accumulated load construct. This means that accumulated load should mediate the relationship between convenience and satisfaction, and since both consequences of inconvenience are captured by accumulated load, no other effect will remain by which convenience can affect satisfaction. Thus, we expect a full mediation effect.

H5: The relationship between convenience and user satisfaction is fully mediated by accumulated load.

Figure 4 illustrates our research model.

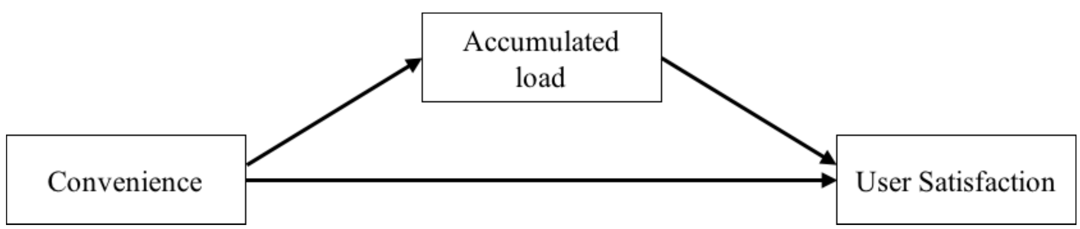


Figure 4- Research Model of Study 2

Method

In the second experiment, we study the effect of convenience on cognitive load, and the effect of the latter on user satisfaction. Similar to Study 1, we chose online grocery to test our research model. The experiment was designed with the cooperation of a well known online grocery retailer in Canada (Sobeys Corporate). This process brought further practical implications to this research by applying the results to the design of an existing online grocery website.

Experimental Design

A single-factor (low/medium/high search convenience) within-subject experiment was designed to test the convenience of an online shopping website. The low convenience condition is represented by a manual product search to find the desired products. Manual search is the basic search function of the website; users have to write product keywords in the website's search box and select their desired products among the items presented on the results page. However, online grocery shopping is a type of IT task that usually includes shopping for several items. Thus, many online groceries have developed a feature on their website that enables users to search for multiple items within a single search query. For instance, Tesco[2], the largest online grocery retailer in the UK, provides its online consumers with the multiple product search feature. This multi-search feature lets users type the keywords of all products in a search box separated by a space (e.g., milk bread orange juice). The website responds by showing a results page for each product respectively. The multi-search function of the online grocery website was used as the medium convenience condition. In order to improve the convenience of the multi-search function, we developed a modified multi-search feature in collaboration with a local online grocer. Following a review of the current multi-search features used by grocery companies across the world, two focus groups provided the designers with the ways to improve the convenience of the existing multi-search feature of the online grocer. More specifically, three features of the function were improved: 1) Accessibility to the multi-search function was improved. 2) Product presentation and comparison were improved. The modified multi-search allowed users to scroll horizontally to compare products in the same category (e.g., chocolate bars) and vertically to see other products they are searching for. 3) Multiple product keywords could be entered in a list (this feature has been implemented in a few online grocery websites such as Sainsbury's[3] and Waitrose[4]). The modified multi-search was used as the third condition, i.e., high convenience. Figure 5 illustrates the difference between the three search functions.

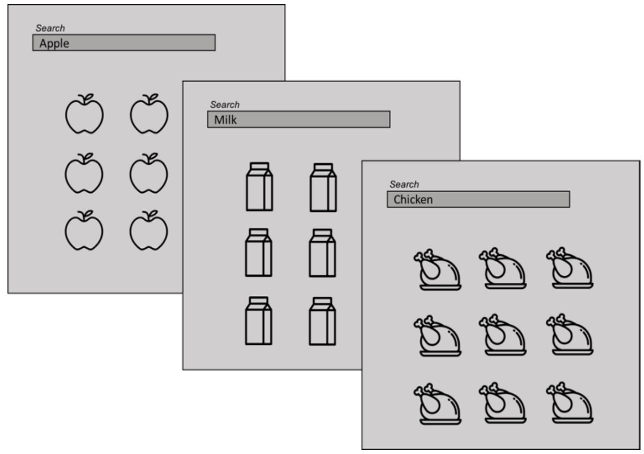
[2] https://www.tesco.com/. The largest grocery retailer in the UK with 28.4% market share

[3] https://www.sainsburys.co.uk/. Sainsbury's is the second largest grocery chain in the United Kingdom

[4] http://www.waitrose.com/. Waitrose is a UK based grocery chain, which has 5.1% share of the UK market

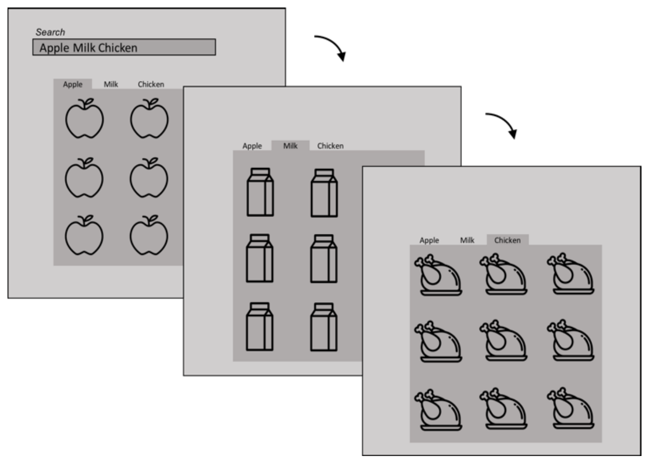
Condition 1 – Low Convenience

Participants had to search manually for multiple products. For each product, they entered the keyword, browsed the results, and selected a product. They did the same steps for other products.



Condition 2 – Medium Convenience

Multiple keywords could be entered in one search inquiry. In the first page, they were presented with the choices of the first product they searched for. Then they could proceed to the next pages to select other products



Condition 3 – High Convenience

Product names could be entered in a list or similar to the old multi-search function. All products were presented in one page. Participants could scroll horizontally to visualize different choices of the same product

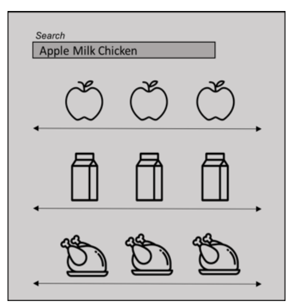


Figure 5- Experimental Condition

Sample and Procedure

Thirty subjects (N=30, 15 males) participated in the experiment. The subjects first performed the low convenience task to understand the website design and structure, and to become familiar with the experimental task. They then performed the two other tasks in random order. All participants were recruited from the university panel and compensated with a $30 gift card for participating in the study. This research was approved by the institution's ethics committee.

Each experimental task included shopping for 10 grocery items. Participants started their task on a recipe page and were asked to search for ingredients and add their product choices to the shopping cart. Participants used a single search function in the first task (Low convenience) and had to search for each item separately. In the Medium convenience condition, participants used the current multi-search function. They could search for multiple items by typing the names of all products in the search box with a space between the names. In the High convenience condition, participants performed the same task using the modified multi-search function. At the end of each task, participants filled out a questionnaire measuring their self-reported cognitive load and satisfaction. The experimental procedure is presented in Figure 6.

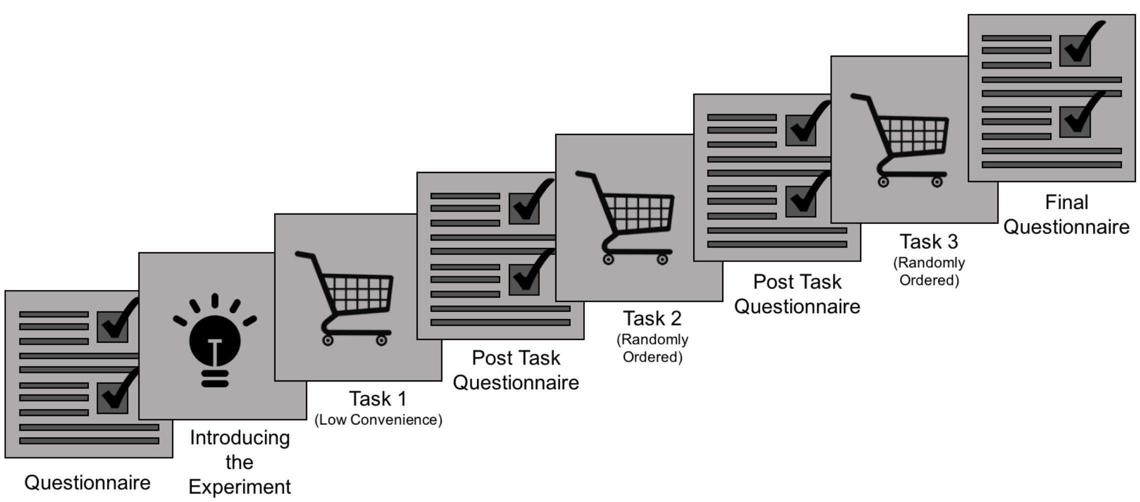


Figure 6- Experimental Procedure of Study 2

A pretest with five participants was performed in order to test measurement tools and the experimental procedure.

Apparatus and Measures

We used the same linear algorithm as Study 1 to measure instantaneous workload. The area under the instantaneous workload curve was used as accumulated load. EEG data was recorded with a 32 electrodes using EGI's (Electrical Geodesics, Inc., Eugene, United States) dense array electroencephalography (dEEG).

To test our manipulation of convenience, the psychometric measure of convenience was measured following each task. The Likert scale developed by Jiang et al., (Jiang et al., 2013) was used. Low values show participants' agreement with the convenience of search tools (All Study 2 measurement scales are presented in Appendix 1). User satisfaction with the shopping experience was measured after finishing each task using a scale developed by Maxham & Netemeyer (Maxham III & Netemeyer, 2002).

We measured four control variables in this study. Need for cognition is defined as "an individual's tendency to engage in and enjoy effortful cognitive endeavors" [59, p. 197]. It is an individual difference construct, which is expected to affect the user's level of mental workload. High need for cognition individuals tend to engage cognitively in a task more than low need for cognition individuals (Cacioppo, Petty, Feinstein, & Jarvis, 1996). This construct was measured with an 18-item scale developed by Cacioppo et al. (Cacioppo et al., 1996). Three types of experience were measured to control for the user's familiarity with the task: 1) Experience with online shopping, 2) Experience with grocery shopping, and 3) Experience with online grocery shopping. Each of the variables was measured with a single item asking how much experience participants have. We expect users who are more experienced with online shopping, grocery shopping, or online grocery shopping to use less time and effort in making their product decision.

Data Processing

EEG data cleaning and processing was performed according to the same steps provided in Study 1 in NeuroRT Studio 2.5.0.3 (Mensia Technologies, Paris, France). This application enables us to create a scenario to process the signal and product the desired output.

Results

Manipulation Check

To ensure that our experimental design was able to manipulate three different levels of search convenience, the participant's perception of search convenience was measured following the conclusion of each shopping task. Reliability analysis shows that the scale is reliable (Cronbach alpha=.88). ANOVA for repeated measures shows that perceived search convenience differs between experimental tasks (M_{Low}=1.94, M_{Medium}=3.69, M_{High}=4.85, $p< .001$). Next, we compared the groups pairwise to make sure there are significant differences between all groups (Low vs. Medium, $p< .01$; Low vs. High, $p< .05$; Medium vs. High, $p<.01$). All pairwise comparisons show significant differences between the three groups in terms of the participants' perception of search convenience. Therefore, our experimental manipulation was satisfactory.

Hypothesis Testing

H3 suggests a relationship between search convenience and the user's accumulated load. We performed panel regression analysis using the Stata software package (StataCorp, Texas, United States), which allows us to account for the dependence of within-subject observations. Table 1 shows that the relationship between convenience and accumulated load is significant ($p < .001$), providing support for Hypothesis 3. Results also show that task order ($p=0.529$) had no effect on the user's accumulated load. No effect was found for Need for cognition ($p=0.695$), Experience with grocery shopping ($p=0.281$), and Experience with online grocery shopping ($p=0.652$); however, Experience with online shopping ($p<0.05$) was found to be a significant control variable.

Table 3- Regression Analysis. DV=Accumulated Load

Accumulated load	Coefficient	Std. Err.	Z	P Value	95% conf. Interval	
Online shopping experience	-0.23	0.10	-2.22	0.027	-0.439	-0.026
Grocery shopping experience	-0.09	0.08	-1.08	0.281	-0.265	0.077
Online grocery shopping experience	0.07	0.16	0.45	0.652	-0.249	0.397
Need for Cognition	0.14	0.36	0.39	0.695	-0.561	0.844
Task order	-0.02	0.04	-0.63	0.529	-0.120	0.062
Convenience	-0.15	0.04	-3.39	**0.001**	-0.249	-0.066
Constant	10.97	0.86	12.71	0.000	9.2795	12.664

Our fourth hypothesis suggests a negative relationship between accumulated load and user satisfaction. Similar to the first hypothesis, we performed a panel regression analysis. As provided in Table 4, results show that the relationship between satisfaction and accumulated load is negative and significant at $p<.05$, which provides support for the Hypothesis 4. Results provide no effect for Order ($p=0.655$), Need for cognition ($p=607$), Experience with online shopping ($p=0.151$), Experience with grocery shopping ($p=0.760$), and Experience with online grocery shopping ($p=0.694$).

Table 4- Regression Analysis. DV=Satisfaction

Satisfaction	Coefficient	Std. Err.	Z	P Value	95% conf. Interval	
Experience/online shopping	-0.33	0.23	-1.44	0.151	-0.791	0.121
Experience/grocery shopping	-0.05	0.18	10.31	0.760	-0.416	0.303
Experience/online grocery shopping	0.14	0.36	0.39	0.694	-0.565	0.850
NFC	0.05	0.10	0.51	0.607	-0.156	0.267
Order	0.09	0.21	0.45	0.655	-0.325	0.517
Accumulated Load	-0.72	0.32	-2.23	**0.026**	-1.363	-0.086
Constant	10.85	3.94	2.75	0.006	3.131	18.582

H5 posits that accumulated load fully mediates the relationship between convenience and satisfaction. We used MLmed (Rockwood & Hayes, 2017) macro in SPSS to estimate the mediation effect. Our data structure is 1-1-1, meaning that all the variables are measured at the lowest level of measurement (i.e. 1 measure per individual per condition). The MLmed Macro estimates the indirect effect and Monte Carlo confidence intervals. The results suggest that there is no mediating effect of accumulated load in the relationship between convenience and satisfaction, thus H5 is not supported. The results are provided in Table 5.

Table 5- Mediation Analysis Result

<table>
<tr><td colspan="7">Mediation Analysis: Indipendent Variable = Conveniece, Dependent Variable = Satisfaction, Mediator = Accumulated load</td></tr>
<tr><td></td><td>Observed Coef.</td><td>SE</td><td>Z</td><td>P>|z|</td><td colspan="2">Monte Carlo Confidence Interval</td></tr>
<tr><td>Indirect effect</td><td>0.006</td><td>0.08</td><td>0.07</td><td>0.93</td><td>-0.156</td><td>0.1719</td></tr>
</table>

Discussion

We designed an experiment to introduce an accumulated load construct as a criterion to evaluate IT artifacts. We used EEG to measure the user's accumulated load, which sums all cognitive load that the user experiences over time. The convenience construct which had been studied before in marketing literature (Jiang et al., 2013; Moeller et al., 2009), was linked to accumulated load because it relates to both dimensions of accumulated load. On the other hand, accumulated load is related to user satisfaction, an already identified business need in an online shopping context. Our results provide support for the causal relationships between accumulated load and convenience, and between accumulated load and satisfaction (H3 and H4)

Our fifth hypothesis, which hypothesized the full mediating effect of accumulated load on the link between convenience and user satisfaction, was not supported. There is a possible explanation for this result. Zhao et al. (Zhao, Lynch Jr, & Chen, 2010) suggest that in conditions where only the direct effect is significant, it is possible that there are other mechanisms for the effect of an independent variable on the dependent variable. It means that convenience has possibly other means of affecting user satisfaction, i.e., other than cognitive load and time. Satisfaction is an affective reaction to an experience with a product or service (Spreng & Mackoy, 1996). It also has a cognitive dimension resulting from the user's appraisal of the experience with the product or service. Consumers feel satisfied insofar as their expectations are confirmed by their use of the system (Bhattacherjee, 2001). One possible explanation of other mechanisms through which convenience influences satisfaction could be the effect of convenience on user expectations. High convenience may induce a high perception of information or system quality which increases the confirmation of the user's expectations prior to using the system, and consequently increase user satisfaction. Although H5 is not supported, it provides an interesting result which can be a motivation to further advance the theoretical foundation of consumer convenience.

1.5 General Discussion and Concluding Remarks

The general goal of this research was to address the potential of NeuroIS to inform design science research. To accomplish this goal, we designed two studies to 1) develop a tool for evaluating an IT artifact, 2) theoretically link it to a design criterion and 3) evaluate an existing IT artifact using the new measure.

In our first study, we measured the user's instantaneous workload using EEG. Then, we extracted three different variables reflecting different aspects of the cognitive load construct: average load, peak load, and accumulated load. These measures had already been conceptualized in the literature (Paas, Tuovinen, et al., 2003; Xie & Salvendy, 2000) but, to the best of our knowledge, had not been empirically tested. Our results show that all workload measures (self-reported, average, peak, accumulated) are sensitive to task difficulty (H1). This effect had been found in prior studies only for one type of cognitive load (Gopher & Braune, 1984). Our results show that only accumulated load was sensitive to the task uncertainty factor (H2). As mentioned, uncertainty impairs information perception, which is the first step in decision making (Xie & Salvendy, 2000). One interpretation of the result is that this effect prolongs the decision making process. Thus, the only measure that could capture its effect is accumulated load.

In our second study, we provided a theoretical explanation for the importance of the accumulated load construct. Our results show that convenience affects accumulated load (H3). Lack of convenience had been associated with time and mental effort in the literature by Jiang et al. (Jiang et al., 2013). Accumulated load however had not been studied in relation to consumer convenience. Our results support prior knowledge that cognitive load is a predictor of user satisfaction (H4).

Our study contributes to theory by differentiating between the four types of cognitive load. The importance of different cognitive load measures had been proposed in the education literature before (Paas, Tuovinen, et al., 2003), but had not been discussed in relation to user-IT interactions. We proposed that these measures reflect different types of cognitive load, therefore the use of them in research or practice depends on the conceptual framework and purpose of the project.

Our study contributes to research by developing a measurement tool to capture various types of cognitive load. To the best of our knowledge, this study is the first that has measured different types of cognitive load and examined the effect of two workload factors (task difficulty and uncertainty) on these four measures. Our results show that although all four measures are labeled as cognitive load measures, they explain different aspects of this multi-faceted construct.

Another contribution of our study is explanation of the relationship between accumulated load and convenience and also accumulated load and user satisfaction. It enriches design theories by introducing a new construct for evaluating the convenience of an information technology. Although time and effort had been proposed as convenience factors (Liang et al., 2015), the relationship between convenience and accumulated load had not been studied before. Furthermore, our research model presents the nomological net in which convenience, accumulated load, and satisfaction exist. It provides an explanation for one of the many mechanisms by which consumers perceive satisfaction.

Our study also has implications for practice by giving designers a validated measure and method to increase user satisfaction and convenience. This method is powerful not only because it is more comprehensive in capturing the cognitive load construct than self-reported measures, but also because of its high temporal resolution, which enables designers to evaluate major and even minor elements of an IT artifact with high precision. For instance, a user's experience with an IT artifact can be broken down to experience with different features of the website. A shopping experience can be divided into the user's primary experience with the home page, then search experience, evaluation experience etc. A user's accumulated cognitive load can be calculated for each of the activities on the website and be used to evaluate the respective design elements.

Another contribution of this study to practice is the introduction of the peak load measure. Our study provides a sigle measurement of peak load which could be used to understand a unique aspect of users experience with IT. Specifically, it can be used to highlight the moments that users have experienced high cognitive load relative to other moments in a user-IT interaction, and to study why such a phenomenon is observed. For instance, in an online shopping task, users engage in various types of activities in order to accomplish the task. Peak load informs us at what moment during the task users have had difficulty in performing the task. Practitioners can then modify the IT artifact components that cause the high cognitive load, and provide users with a more satisfying shopping experience.

As with any other research endeavor, this research has certain limitations that need to be mentioned. Although neuroscience tools enable us to measure a wide range of cognitive variables that are hard to capture using conventional methods, there is a need to replicate studies using these neurophysiological measures across different contexts in order to improve their generalizability. Thus, we believe that more research is needed to test our method and measures on various IT artifacts in order to further validate them.

There are a number of avenues for advancing this research. First, we believe that other measures extracted from instantaneous workload can be interesting for certain purposes, especially peak load which provides unique information about the exact moments that users are overloaded with an event or feature during their transaction with an IT artifact. Second, experimenting the effect of high peak load versus high accumulated load to see if a very high peak load experienced by a user over a short period (exceeding a cognitive capacity red line) will have a larger impact on the user than experiencing the same amount of accumulated load but over a longer time period. This can help us to have a better understanding of how cognitive load affects user satisfaction. Third, accumulated load can be possibly linked to other design-related criteria such as negative feelings, fatigue, and frustration. Finally, as suggested by prior studies (de Guinea et al., 2014), there could be a difference between implicit (automatic or unconscious) and explicit (self-reported) type of constructs in terms of their effect on user behavior since they represent different aspects of the cognitive load construct. It is possible that the research model of our second study could be revised, and explicit cognitive load be added to it. For instance, it could be interesting to test if explicit cognitive load mediates the relationship between convenience and satisfaction or if it can be used to improve the research model proposed in our second study.

References

Aljukhadar, M., Senecal, S., & Daoust, C.-E. (2012). Using recommendation agents to cope with information overload. International Journal of Electronic Commerce, 17(2), 41-70.

Baddeley, A. (2007). Working memory, thought, and action (Vol. 45): OUP Oxford.

Berry, L. L., Seiders, K., & Grewal, D. (2002). Understanding service convenience. Journal of marketing, 66(3), 1-17.

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

Briggs, R. O., De Vreede, G.-J., & Nunamaker Jr, J. F. (2003). Collaboration engineering with ThinkLets to pursue sustained success with group support systems. Journal of management information systems, 19(4), 31-64.

Brocke, J. V., Riedl, R., & Léger, P.-M. (2013). Application strategies for neuroscience in information systems design science research. Journal of Computer Information Systems, 53(3), 1-13.

Brouwer, A.-M., Zander, T. O., van Erp, J. B., Korteling, J. E., & Bronkhorst, A. W. (2015). Using neurophysiological signals that reflect cognitive or affective state: six recommendations to avoid common pitfalls. Frontiers in neuroscience, 9, 136.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. Psychological bulletin, 119(2), 197.

Cameron, A.-F. (2007). Juggling multiple conversations with communication technology: Towards a theory of multi-communicating impacts in the workplace: Queen's University.

Charland, P., Léger, P.-M., Sénécal, S., Courtemanche, F., Mercier, J., Skelling, Y., & Labonté-Lemoyne, E. (2015). Assessing the multiple dimensions of engagement to characterize learning: A neurophysiological perspective. Journal of visualized experiments: JoVE(101).

Chintagunta, P. K., Chu, J., & Cebollada, J. (2012). Quantifying transaction costs in online/off-line grocery channel choice. Marketing Science, 31(1), 96-114.

Colle, H. A., & Reid, G. B. (1999). Double trade-off curves with different cognitive processing combinations: Testing the cancellation axiom of mental workload measurement theory. Human factors, 41(1), 35-50.

Colle, H. A., & Reid, G. B. (2005). Estimating a mental workload redline in a simulated air-to-ground combat mission. The International Journal of Aviation Psychology, 15(4), 303-319.

Colwell, S. R., Aung, M., Kanetkar, V., & Holden, A. L. (2008). Toward a measure of service convenience: multiple-item scale development and empirical test. Journal of Services Marketing, 22(2), 160-169.

Coyne, J. T., Baldwin, C., Cole, A., Sibley, C., & Roberts, D. M. (2009). Applying real time physiological measures of cognitive load to improve training. Paper presented at the International Conference on Foundations of Augmented Cognition.

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

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

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. Instructional science, 38(2), 105-134.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. Journal of educational psychology, 100(1), 223.

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

DeStefano, D., & LeFevre, J.-A. (2007). Cognitive load in hypertext reading: A review. Computers in human behavior, 23(3), 1616-1641.

Dimoka, A., Davis, F. D., Gupta, A., Pavlou, P. A., Banker, R. D., Dennis, A. R., . . . Gefen, D. (2012). On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS. Mis Quarterly, 679-702.

Garbarino, E. C., & Edell, J. A. (1997). Cognitive effort, affect, and choice. Journal of consumer research, 24(2), 147-158.

Garrett, D., Peterson, D. A., Anderson, C. W., & Thaut, M. H. (2003). Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. IEEE Transactions on neural systems and rehabilitation engineering, 11(2), 141-144.

Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? Human factors, 26(5), 519-532.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept.

Gwizdka, J. (2010). Distribution of cognitive load in web search. Journal of the Association for Information Science and Technology, 61(11), 2167-2187.

Hevner, A. R. (2007). A three cycle view of design science research. Scandinavian journal of information systems, 19(2), 4.

Hoque, A. Y., & Lohse, G. L. (1999). An information search cost perspective for designing interfaces for electronic commerce. Journal of marketing research, 387-394.

Jiang, L., Yang, Z., & Jun, M. (2013). Measuring consumer perceptions of online shopping convenience. Journal of Service Management, 24(2), 191-214.

Kim, C., Mirusmonov, M., & Lee, I. (2010). An empirical examination of factors influencing the intention to use mobile payment. Computers in human behavior, 26(3), 310-322.

Liang, H., Peng, Z., Xue, Y., Guo, X., & Wang, N. (2015). Employees' exploration of complex systems: an integrative view. Journal of management information systems, 32(1), 322-357.

Libenson, M. H. (2012). Practical Approach to Electroencephalography E-Book: Elsevier Health Sciences.

lo Storto, C. (2013). Evaluating ecommerce websites cognitive efficiency: An integrative framework based on data envelopment analysis. Applied ergonomics, 44(6), 1004-1014.

Maxham III, J. G., & Netemeyer, R. G. (2002). A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. Journal of marketing, 66(4), 57-71.

Mikulka, P. J., Scerbo, M. W., & Freeman, F. G. (2002). Effects of a biocybernetic system on vigilance performance. Human factors, 44(4), 654-664.

Minas, R. K., Potter, R. F., Dennis, A. R., Bartelt, V., & Bae, S. (2014). Putting on the thinking cap: using NeuroIS to understand information processing biases in virtual teams. Journal of management information systems, 30(4), 49-82.

Mizuno, K., Tanaka, M., Yamaguti, K., Kajimoto, O., Kuratsune, H., & Watanabe, Y. (2011). Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. Behavioral and brain functions, 7(1), 17.

Moeller, S., Fassnacht, M., & Ettinger, A. (2009). Retaining customers with shopping convenience. Journal of Relationship Marketing, 8(4), 313-329.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology, 48(2), 229-240.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. Educational psychologist, 38(1), 1-4.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. Educational Psychology Review, 24(1), 27-45.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. Educational psychologist, 38(1), 63-71.

Qiu, L., & Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. Journal of management information systems, 25(4), 145-182.

Rockwood, N. J., & Hayes, A. F. (2017). MLmed: An SPSS macro for multilevel mediation and conditional process analysis. Paper presented at the Poster presented at the annual meeting of the Association of Psychological Science (APS), Boston, MA.

Ryu, K., & Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. International Journal of Industrial Ergonomics, 35(11), 991-1009.

Seiders, K., Berry, L. L., & Gresham, L. G. (2000). Attention, retailers! How convenient is your convenience strategy? Sloan Management Review, 41(3), 79.

Seiders, K., Voss, G. B., Godfrey, A. L., & Grewal, D. (2007). SERVCON: development and validation of a multidimensional service convenience scale. Journal of the academy of marketing science, 35(1), 144-156.

Sénécal, S., Léger, P.-M., Riedl, R., & Davis, F. D. (2018). How Product Decision Characteristics Interact to Influence Cognitive Load: An Exploratory Study Information Systems and Neuroscience (pp. 55-63): Springer.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. Cerebral Cortex, 6(1), 11-20.

Spreng, R. A., & Mackoy, R. D. (1996). An empirical examination of a model of perceived service quality and satisfaction. Journal of retailing, 72(2), 201-214.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. Learning and instruction, 4(4), 295-312.

Todd, P., & Benbasat, I. (1992). The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. Mis Quarterly, 373-393.

Todd, P., & Benbasat, I. (1999). Evaluating the impact of DSS, cognitive effort, and incentives on strategy selection. Information systems research, 10(4), 356-374.

Turel, O., & Qahri-Saremi, H. (2016). Problematic use of social networking sites: antecedents and consequence from a dual-system theory perspective. Journal of management information systems, 33(4), 1087-1116.

Venables, L., & Fairclough, S. H. (2009). The influence of performance feedback on goal-setting and mental effort regulation. Motivation and Emotion, 33(1), 63-74.

vom Brocke, J., & Liang, T.-P. (2014). Guidelines for neuroscience studies in information systems research. Journal of management information systems, 30(4), 211-234.

Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical issues in ergonomics science, 3(2), 159-177.

Wilson, T. D., Ford, N., Foster, A., & Spink, A. (2000). Uncertainty and its correlates. The New Review of Information Behaviour Research, 1(December), 69-84.

Xie, B., & Salvendy, G. (2000). Prediction of mental workload in single and multiple tasks environments. International journal of cognitive ergonomics, 4(3), 213-242.

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. Journal of consumer research, 37(2), 197-206.

Appendix I – Measures

Convenience

The web site is user-friendly for making purchases.

The web site is easy to understand and navigate.

I am able to find desired products quickly.

The product classification is intuitive and easy to follow.

Satisfaction

I am satisfied with my overall experience with …

As a whole, I am not satisfied with …

How satisfied are you overall with the quality of …

Cognitive Load

I spent a lot of mental effort doing the task

The task required a great deal of mental effort

The task required a great deal of concentration

The task did not require much mental effort

I had to work mentally to do the task

Need for Cognition

I would prefer complex to simple problems.

I like to have the responsibility of handling a situation that requires a lot of thinking.

Thinking is not my idea of fun.

I would rather do something that requires little thought than something that is sure to challenge my thinking abilities?

I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something."

I find satisfaction in deliberating hard and for long hours.

I only think as hard as 1 have to.

I prefer to think about small, daily projects to long-term ones?

I like tasks that require little thought once I've learned them?

The idea of relying on thought to make my way to the top appeals to me.

I really enjoy a task that involves coming up with new solutions to problems.

Learning new ways to think doesn't excite me very much?

I prefer my life to be filled with puzzles that I must solve.

The notion of thinking abstractly is appealing to me.

I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.

1 feel relief rather than satisfaction after completing a task that required a lot of mental effort? It's enough for me that something gets the job done; I don't care how or why it works?

I usually end up deliberating about issues even when they do not affect me personally.

Chapter 3-
Essay II - Investigating the Effect of the Match Between Product Sorting and User Goals on Cognitive Load and Shopping Performance

Abstract

This study investigates the contingent effect of product sorting and users' goal on cognitive load and performance. Our research model posits that product sorting will help users by reducing their cognitive load level provided that the sorting type matches users' criteria for choosing products. A 2 X 2 (Product Sorting X User goal) within subject was designed to test the hypotheses. Multiple measures of cognitive load was obtained using Event-Related-Potentials and frequency power analysis methods. Our results show that the match between user goal and product sorting reduce user cognitive load. All six measures of cognitive load were influenced by the match variable. Our analysis also shows that cognitive load negatively predicts user performance as hypothesized. Time to perform the shopping task (performance) is affected by five measures of cognitive load. Contributions for theory and practice are discussed.

2.1 Introduction

Information display is an important element in online shopping, because on the one hand it can be flexible and be designed in very different ways (West et al., 1999) and on the other hand it affects consumers' decisions and behaviors (Cai & Xu, 2008). Product result pages often provide different types of item sorting (e.g., alphabetical, price), which is an instance of changing information display. Sorting may act as a decision support tool for consumers (Sharkey et al., 2009). It changes information display in order to help consumers find their desired products (Ariely, 2000). Similar to other decision support tools, sorting can be used to improve the user's decision making. However, it is not clear under what conditions various types of sorting may decrease or increase the user's cognitive effort during the decision-making process.

Minimizing cognitive load is important for users in online shopping, and it is even more important when it comes to shopping for low value products. There are two ways that sorting may contribute to the enhancement of consumer decision making: 1- Improving decision quality and 2- Saving the user's cognitive effort (Todd & Benbasat, 1992). Researchers have found that the trade-off between the two factors (i.e., maximizing accuracy and minimizing effort) depends on the task and the context (Payne et al., 1988, Beach & Mitchell, 1978). For instance, it is more likely that a typical consumer puts more effort into obtaining accuracy when buying an apartment than grocery items. We use the same logic to explain how sorting may affect the user's level of mental effort in shopping for low price goods. In shopping for these goods, consumers are expected to show lower decision accuracy in favor of minimizing their mental effort. It is also important to understand how sorting affects the user's cognitive load since it is a predictor of user satisfaction and performance in online shopping (lo Storto, 2013). Product sorting could increase user satisfaction and performance; however, these relationships need to be explained theoretically with respect to the user's goal and be tested empirically.

Previous research suggests that an appropriate information sequence can result in an easier decision making process (Schkade & Kleinmuntz, 1994). This effect is contingent upon the alignment of the information sequence with what the user is looking for. Information sequence can facilitate the decision process if it increases the user's accessibility to the right information. For instance, if a consumer is looking for a specific price or brand name, sorting products based on price or alphabetically can help make their product decision,

In this research, we design a two factor within-subject experiment (Product sorting X Users goal) and hypothesize that if product sorting matches the user's goal, it decreases user cognitive load. We also expect a negative relationship between cognitive load and shopping performance, because the less cognitive load consumed in a task greater is the efficiency in performing it.

Our study contributes to theory by showing how the presentation sequence of information (i.e., sorting) affects user cognitive load in a decision making process. We model this effect as a relationship between a fit construct (i.e., match between product sorting and user goal) and cognitive load. It contributes to methodology by providing a new experimental design to study user behavior using an Event-Related-Potential (ERP) method in a user-IT transaction. Unlike traditional ERP research, we use one of the users's activities during the shopping session to generate ERPs. It also has practical implications since it shows how various types of product sorting, currently used in shopping websites, can reduce the user's mental workload. Our study also shows that not offering users the appropriate product sorting variables (e.g., date, price, and alphabet) reduces their shopping performance.

2.2 Literature Review

Researchers propose that users seek to maximize their decision quality and minimize the cognitive effort exerted during this process (Todd & Benbasat, 1992). Consumers, depending on the context, find a trade-off between the two factors, and make their product decision. However, in some contexts, the importance of decision quality is lower compared to minimal cognitive effort (Bettman et al., 1998). For instance, compare shopping for an apartment and grocery items. Any mistake in the former decision may have serious effects on the user's life and be hard to recover from, whereas in the latter the user is less sensitive to the accuracy of the decision since in the worst case it will be easy to buy another product. Even generally, cognitive effort is proposed to bear more weight than accuracy (Todd & Benbasat, 1992). The reason behind this phenomenon is that the feedback from effort expenditure is immediate in comparison to feedback from accuracy, which takes more time to operate (Kleinmuntz & Schkade, 1993). Therefore, accuracy is sacrificed in favor of saving cognitive load, and the intensity of such sacrifice depends on the decision making task and context.

Cognitive effort is an important factor in explaining human decision making. It is considered as the cost of decision making for users (Todd & Benbasat, 1992). Cognitive load is defined as the set of mental resources used by people to encode, activate, store, and manipulate information while they perform a task (DeStefano & LeFevre, 2007). A key to understanding cognitive load and its effect on human behavior is that these mental resources are limited (Wickens, 2002). Therefore, efficient use of working memory is a key factor in preventing users from overload situations and providing them with a satisfying shopping experience (Aljukhadar et al., 2012; lo Storto, 2013). In an online shopping session, any website element that fails to provide users with the critical information needed for making product decisions reduces cognitive efficiency of the website (lo Storto, 2013). This failure could be either not providing necessary information or presenting redundant information to users. Poor design of shopping websites results in consumers needing to devote more working memory resources (e.g., attentional capacity of working memory) to accomplish the shopping task. A number of factors affect the user's cognitive workload, among them are different ways of presenting information, which include form (numerical, pictorial, verbal), organization (table, matrix, list, paragraph, hierarchical cluster), and sequence (random, ascending or descending on an attribute value, alphabetical, chronological) (Kleinmuntz & Schkade, 1993, Todd & Benbasat, 1992).

In the online shopping environment, product sorting is a form of information sequence modification. It presents information to the consumer in a new way to assist in making a product decision. Sorting arranges products based on a specific attribute and helps consumers to narrow down their consideration set (Sharkey et al., 2009). In this sense, sorting can be considered as a simple decision support tool because one of the functions of decision support systems is screening and sorting alternatives (Van der Heijden, 2006). It supports consumer decision making by determining the relative utility of alternatives (Häubl & Trifts, 2000). It contributes to the minimization of the consumer's' mental workload. Consumers will be able to screen alternatives and more easitly reduce the universal set to a consideration set.

Online shopping tasks can be classified into two general groups based on the consumer's goals: searching versus browsing tasks (Carmel et al., 1992). In searching tasks, the consumers' objectives and criteria are clear (Hong et al., 2004) since they know in advance what product they are looking for. For instance, they may know the brand name of the product. In contrast, consumers engaged in a browsing task have no specific criteria (Hong et al., 2004). For instance, they may simply have the intention of buying a TV, however, this does not mean that they do not have any criteria when it comes to actually purchasing the TV. In the current study, we are focusing on search tasks.

2.3 Hypotheses Development

We theorize that product sorting affects the user's mental workload depending on the user's goal. Product sorting (i.e., listing products based on the sequence of their values) can decrease the user's cognitive effort if it is aligned with the user's goal. Cognitive load also negatively influences user performance. Figure 7 illustrates our research model.

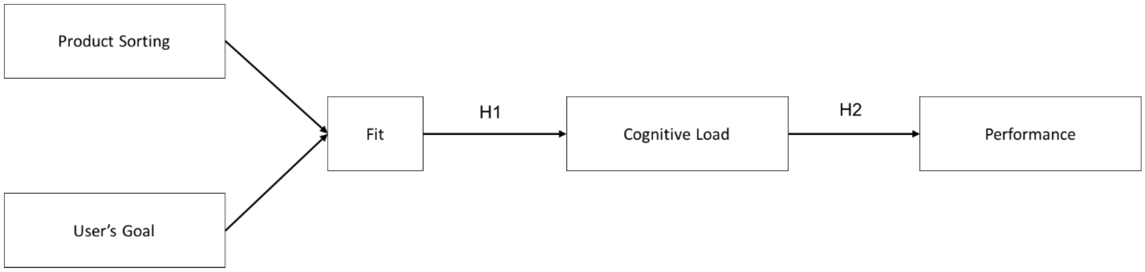


Figure 7- Research Model

We model the contingent effect of product sorting and the user's goal as a fit construct. Venkatraman (1989) proposed six conceptualizations of a fit construct including fit as matching. Matching represents the fit between two constructs without reference to a criterion construct, however its effect on different sets of criterion variables can be investigated. Therefore, we have either "match" or "mismatch" conditions between the user's goal and product sorting. In match conditions, product sorting assists users to find the target product whereas in mismatch conditions, there is no complementarity between the two variables, and users experience more mental workload to find the target product compared to a match condition. Table 6 shows match (1 and 4) and mismatch conditions (2 and 3).

Table 6- Match Table

		Goal	
		Price	Name
Sorting	Price	1	2
	Brand	3	4

Light Grey: Match Conditions
Dark Grey: Mismatch Condition

As stated, product sorting as a decision support tool helps consumers to make their product decision more efficiently (Cai & Xu, 2008). Consumers will be able to remove a number of items from the universal set without devoting attentional capacity of their working memory. We argue that if a user's goal matches the product sorting on a website, it reduces the user's mental workload. For instance, users who are looking for the cheapest product will be supported by sorting products based on price. Sorting based on brand name will not be useful to them because they need to screen all the product prices.

The reason why a mismatch condition imposes high load on the user's working memory is also explained by eye movements. Eye movement is linked to the attentional focus of users (Postle et al., 2006; Baddeley, 2007). As the eyes move over a screen and explore different objects, the user switches attention from one object to another. This consumes the available attentional resources of working memory, which explains why the user's short term memory performance declines as eye movements increase (Hale el al., 1996). The mismatch condition forces the user to explore objects one by one until finding the target product. It increases eye movements and consequently imposes more load on the user's working memory. Based on the above, our hypothesis is:

H1: Users experience less mental workload when their goal matches product sorting compared to when it does not match.

Cognitive load explains the amount of working memory resources used by consumers in a task. Therefore, saving cognitive effort is one of the criteria for improving decision making in online shopping (Todd & Benbasat, 1992). As the level of cognitive load required for the same task decreases, users are able to perform the task more easily and faster (Hong et al., 2004). Users are obliged to create their consideration set by manually screening every product presented to them and remove those which do not meet their criteria. Thus, we hypothesize the negative relationship between mental workload and task performance.

H2: Mental workload negatively influences task performance.

2.4 Methodology

Experimental Design

A 2 (Product sorting) X 2 (Users' goal) within-subject experiment was designed to test the research hypotheses. Two types of product sorting (Price and Alphabetical) were manipulated in a search result page with ten products (in two rows) as shown in Figure 8. Ten experimental tasks were designed for each condition with different products, which means each participant performed 40 tasks in total. There was no time limit for performing each task, and after selecting the product, subjects were automatically presented with the next task. The product results pages presented to the participants were screenshots whose design was based on a popular regional online grocery website. Participants were asked to select a product based on either a specific price or a brand name, while the products were sorted either by price or alphabetically. The experiment was designed using E-prime 3.0 software (Psychology Software Tools, Pittsburgh, PA). The product brand names were fictitious and unknown to participants. We used product pictures from a real online grocery website.

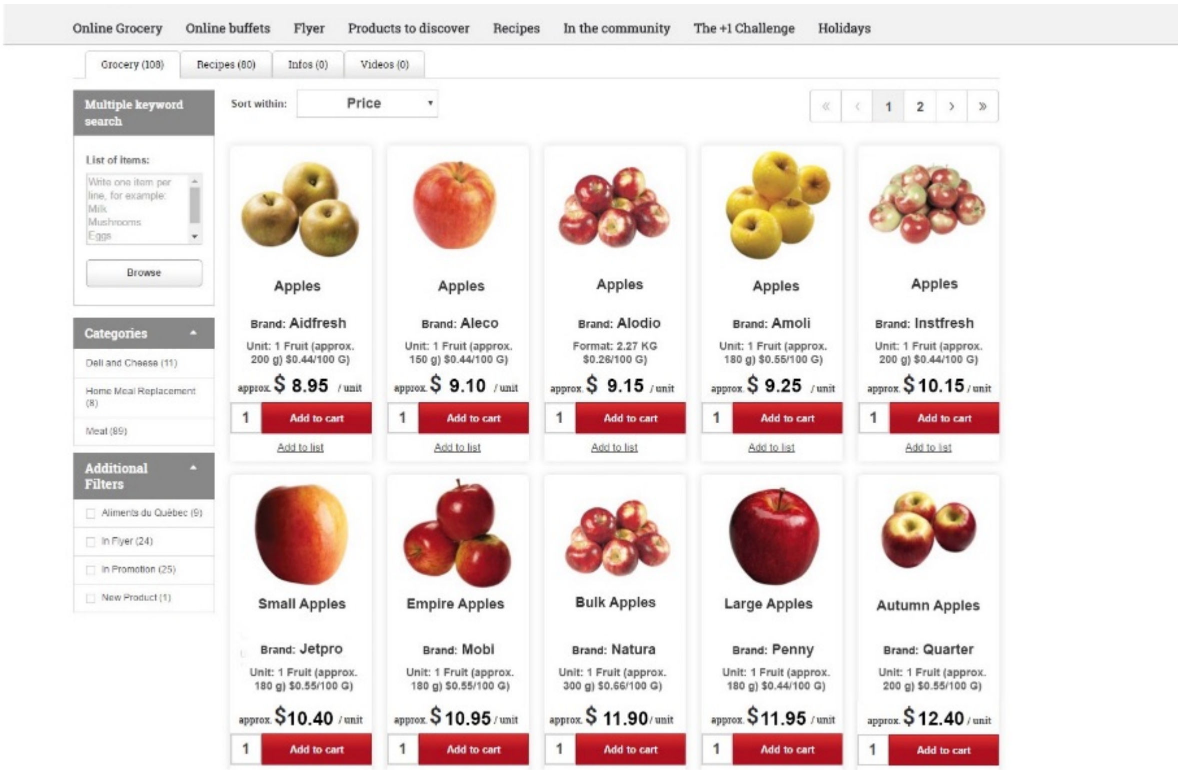


Figure 8- A Sample Result Page with Price Sorting

Before the presentation of each screenshot, an instruction appeared on the screen asking the participant to select a product either based on the price (Goal=price) or a brand name (Goal=brand name). An instruction example for each goal type is provided in Table 7.

Table 7- Instruction Examples

Goal	Instruction Examples
Price	Click on the *cheapest* product / Click on the product with *16.99* price
Brand name	Click on the product with *Jetpro* brand

Sample and Procedure

Twenty subjects (N=20; 50% female) were recruited from a university panel to participate in the experiment. They were compensated with a $30 gift card. Participants were first greeted and then asked to read and sign the consent forms. Then, a 64-channel Brainvision EEG headset (Brain Products GmbH, Gilching, Germany) was placed on participants' head and impedance was tested to ensure the quality of EEG data. Then, the experiment started and participants followed the experimental protocol according to Figure 9. They first filled out the questionnaire, then read the experiment introduction message. A sample page was shown to participants to familiarize them with the experimental task. As part of the introduction, three pages were presented to subjects in which the sort box, price tag, and brand names were highlighted respectively to make sure they knew where to find them on the page. Participants were instructed to click on the image of the target product. Then, subjects were asked to perform a practice task and ask any question about the experiment. Finally, they started performing the forty tasks, which were randomized. This study was approved by the ethics committee of our institution.

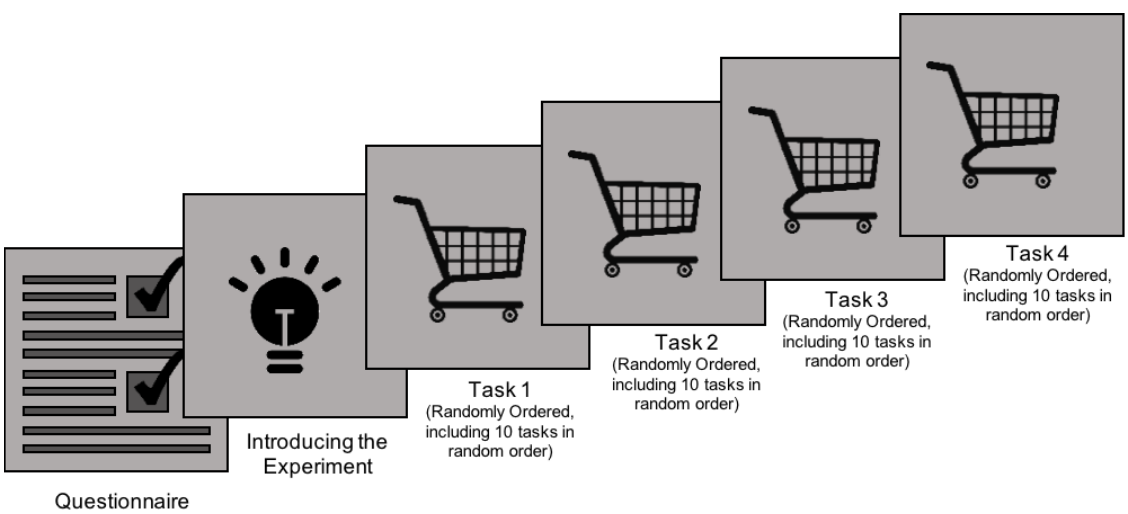


Figure 9- Experiment Procedure

Measurement

Cognitive Load

We used EEG to measure cognitive load. More precisely, we used two methods to capture the user's cognitive load: 1- Event-related Potential (ERP) and 2- Fast Fournier Transform (FFT) analysis. The ERP method is a measure of cognitive load approximately at the time of the event presentation, while the FFT analysis provides the overall amount of cognitive load for the specified time window. Although the ERP gives an instantaneous measure of user workload at the moment of event presentation, given the accumulative nature of cognitive load, it will be affected by the task requirements in the preceding moments. Therefore, we expect that cognitive load measures obtained both by ERP and FFT analysis will be affected by the experimental manipulation.

1. Event-Related-Potential (ERP)

The ERP method is based on EEG (Luck, 2012). Generally, EEG measures the activity of a large group of neurons firing at the same time, and therefore, it is difficult to separate a specific cognitive process associated with that neural activity (Müller-Putz et al., 2015). The ERP method overcomes this problem by presenting stimuli several times and measuring the user's response to them. This would cancel the neural activities unrelated to experimental manipulation (Luck, 2012). Thus, it is crucial to have the exact timing of stimulus presentation to measure the neural activities in phase with it. In classical ERP studies, a task-relevant stimulus such as an oddball task is used to elicit ERPs. This method works well provided that the stimulus presented to the user is phased locked with the investigated cognitive processes (Léger et al., 2014), otherwise the richness of data will be reduced by averaging the epochs and losing the user's real neural response to the stimulus. In this experiment, we use two events to create ERPs: i) the participant's mouse click time stamps and ii) the image presentation time stamps (i.e. the moment that the screenshot of the product result page is presented to the participant). A description of each event is provided in Table 8. The exact time that participants click on the target product is when they have made their decision. The phenomenon of interest in our experiment (i.e. user's' cognitive load at the time of decision making) justifies the selection of this event because we would like to be temporally as close as possible to the moment that users make their product decision. Thus, we generate the ERPs based on the time stamps of the user's decision. We also use the image presentation time to create the ERPs. Subjects performed 10 experimental tasks of each condition sequentially to maintain their cognitive load at the manipulated level (low or high) for the entire duration of the condition. Therefore, ERP measures are affected by the task requirements.

Table 8- Description of image presentation and mouse click events

Event Name	Description
Image Presentation	Right after the instruction page, the screenshot of product results page is presented. The moment that the screenshot is projected on the screen is used as the image presentation event.
Mouse Clicks	The moment that user clicks on the target product in the results page is marked as the mouse click event.

Generally, ERPs consist of a number of important components. These components are found to be sensitive to different cognitive, emotional, and behavioral variables. An ERP sample is illustrated in Figure 10. Components names represent both the polarity and approximate latency of the element. For instance, N100 is a negative peak, which occurs approximately 100 ms after stimulus presentation.

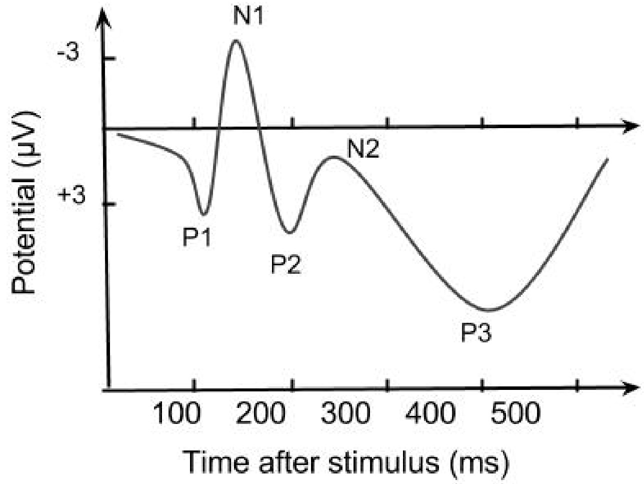


Figure 10- Event Related Potential Components

In this experiment we use the P3 component to measure the user's mental workload. This component is a positive peak, which can be observed approximately 300 ms after the event presentation (i.e., stimulus). It has been found that the amplitude and latency of P3 are sensitive to the user's cognitive load (Murata et al., 2005, Uetake & Murata, 2000). Research shows that P3 amplitude is negatively related to the user's level of cognitive processing load (Ullsperger et al., 1988). Changes in P3 amplitude reflect the distribution of processing resources (Isreal et al., 1980). For instance, if the user performs two concurrent tasks, meaning that cognitive resources are being distributed to meet the requirements of both tasks, the P3 amplitude decreases or disappears (Wickens et al., 1977). Furthermore, P3 amplitude is affected by the difficulty of the ongoing task. For instance, research by Ullsperger et al. (2001) shows that the P3 amplitude drops when users perform difficult arithmetic tasks. P3 latency also increases with the task difficulty (Ullsperger et al., 1986). The latency represents the timing of mental processes, which in a difficult task will take a longer to appear in the EEG (Käthner et al., 2014). Therefore, consistent with the abovementioned study, we use P3 amplitude and latency at thePz electrode to measure the user's' cognitive load associated with making product decisions.

To generate ERP waves, one needs to average the user's' cognitive response to an experimental event. Such event needs to be discrete and also temporally relevant for measuring the phenomenon of interest (Kramer, 1991). The following explains why both the user's mouse click on the target product and the image presentation moment meet both criteria and are relevant for measuring their cognitive load. First, both events are discrete events within the experimental condition. They can be identified precisely as opposed to other moments during the shopping task. Second, the P3 component of mouse clicks and image presentation events reflect the user's cognitive load during the past few seconds of the event. We explain this based on the nature of cognitive load and the literature. Second, cognitive load is cumulative in nature (Baddeley, 2007) and not necessarily the user's instantaneous reaction to a stimulus. It means that the user's cognitive load at each moment is a function of task requirements in the past. Other experiments in the literature have measured the user's cognitive load on a task using the P3 component of their response to an event during the primary task (Kramer, 1991). These events might or might not be relevant to the primary task (Kramer, 1991), but since they happen within the experimental condition, the P3 component reflects the amount of accumulated cognitive load the user experiences at the moment. Both mouse clicks and image presentations are events within the experimental condition (match or mismatch), which reflect the user's level of cognitive load. In each condition, users click 10 times on the target product, therefore the ERP is measured 10 times using this event for each condition. The images are also presented to users 10 times, however we exclude the first task in each condition because the user's cognitive load has not yet been affected by the condition when they are presented with this first task. Therefore, 9 events of this type are used to generate another set of ERPs.

2. Fast Fourier Transform (FFT) Analysis

The EEG signal can be decomposed into a number of frequency band oscillations including delta (0 to < 4 Hz), theta (4 to < 8Hz), alpha (8 to 13 Hz), and beta (>13 Hz). Research has shown the sensitivity of these oscillations to different cognitive and emotional states of individuals. For instance, alpha and theta band EEG has been found sensitive to the user's workload level (Gevins et al., 1998; Gevins & Smith, 2003). It is suggested that as workload increases alpha band power decreases and theta band power increases (Kramer, 1991; Gevins et al., 1998). The FFT measures of workload are able to capture the total amount of cognitive load experienced by users during a task performance period (Mirhoseini et al., 2017). We use Fast Fourier Transform to calculate the power spectral density of EEG and measure the changes in frequency characteristics. We calculate the average power of both alpha and theta bands for the last second of each task. The average of powers across experimental tasks is used as a measure of cognitive workload for each condition. We expect alpha band power to increase for match conditions compared to mismatch conditions (i.e., negatively associated with cognitive load) and theta band power to decrease for match conditions compared to mismatch conditions (i.e., it is positively associated with cognitive load).

Performance

The measure of shopping performance was task duration (i.e., find the focal product), but for the correct responses only. The incorrect responses were removed from the ERP sample. For instance, if there was one wrong answer, then the ERP was calculated based on the other 9 responses out of ten. Instead of including these tasks with a 0 performance measure, we removed them because the cognitive load measure of wrong responses is likely to be affected by the other tasks in the same condition i.e. including them in the analysis with a performance of 0 for wrong answers would be inaccurate. Task duration represents how efficiently users were able to correctly locate the target product. As stated before, in each experimental condition, users performed 10 tasks. Thus, we have a maximum of 10 performance measures for each condition. We used the average duration time of successful tasks as the task performance measure for each condition.

Data Analysis

To process the EEG data, we used Brainvision Analyzer and MATLAB software. EEG raw data was filtered using a FIR filter of order 96 between 0.1 and 30 Hz (Zeyl et al., 2016). For the MATLAB code, please see Appendix 1. The EEG was then re-referenced to the average of all electrodes. Independent component analysis (ICA) was performed to identify the bad components such as eye-blinks and muscle movements (Jung et al., 1998). The bad components were removed using the ADJUST plugin in EEGLAB toolbox (Mognon et al., 2011). Then, the signal was reconstructed in the time domain using inverse ICA. The signal was then segmented with respect to the mouse click time stamps between -200 and 800 ms of the events (Kramer, 1991). As explained, each participant performed 10 product choice tasks per condition, giving 10 segments per condition per participant for mouse click events and 9 segments per condition per participant for image presentation events. An artifact rejection transformation was applied to remove segments where the EEG amplitude was abnormal. There were also a few instances where few segments were removed from the data because participants made the wrong product choice (for more information about segmentation please see Appendix 1).

Then, segments were averaged for each condition, and peak and latency of the P3 component were extracted for the Pz electrode (Murata et al., 2005). Peak and latency are two indicators of the participant's cognitive load during the task performance.

Table 9 provides the average of the P3 amplitude and latency for each experimental condition for both mouse click and image presentation events.

Table 9- The average of P3 amplitude and latency per experimental condition for both events

Goal	Sorting	Fit	P3 Amplitude (μv) (Mouse Click)	P3 Latency (ms) (Mouse Click)	P3 Amplitude (μv) (Image Presentation)	P3 Latency (ms) (Mouse Click)
Price	Price	Match	24.74	423.2	31.20	366.0
Brand name	Price	Mismatch	14.50	438.5	27.68	427.5
Price	Brand Alphabet	Mismatch	15.75	450.1	27.14	422.8
Brand name	Brand Alphabet	Match	21.15	397.0	32.37	403.2

To extract alpha and theta band powers, Fast Fourier Transforms (FFT) in MATLAB were performed for the last second of each task. Then, the power was calculated according to the following formula:

$$P = \frac{1}{N}\sum_{K=0}^{N-1} |X[k]|^2,$$

where X denotes the FFT representation of the signal and N is the length of the signal (Oppenheim et al., 1978). Then, the power values were averaged for each condition. Table 10 shows the average of alpha and theta band powers for each experimental condition.

Table 10- The average of Alpha and Theta band powers per experimental condition

Goal	Sorting	Fit	Alpha Power	Theta Power
Price	Price	Match	5.21	4.19
Brand name	Price	Mismatch	4.77	4.79
Price	Brand Alphabet	Mismatch	4.42	4.11
Brand name	Brand Alphabet	Match	5.44	4.12

To run regression analysis, the independent variables and the fit constructs were coded. A matching table was created which reflects the fit between goal and sorting variables. The value of the fit variable was equal to 1 for match conditions and 0 for mismatch conditions. Since each of the Sorting and Goal variables have only two levels, they are coded as a binary variable (Table 11).

Table 11- Binary codes of the Sorting and Goal variables

Variable = Sorting	
Code	Description
1	Product are sorted based on price
2	Product are sorted based on brand alphabet
Variable = Sorting	
Code	Description
1	Users are asked to chose a product based on a specific price
2	Users are asked to chose a product based on a specific brand name

2.5 Results

According to H1, we expect P3 amplitude to be higher for match conditions compared to mismatch conditions, and P3 latency to be be less for the match conditions compared to mismatch conditions. Since this experiment has a within-subject design, a panel regression analysis was done to test the effect of match on the latency and amplitude of the P3 component. Four regression models were tested using the Xtreg command in the Stata software package (StataCorp, Texas, United States) for each event: Two models with P3 amplitude and P3 latency of the mouse click events as the dependent variable and two models with P3 amplitude and P3 latency of the image presentation events as the dependent variable.

The first two models show that P3 amplitude of mouse click events was affected by the match between the product sorting variable and user's goal. Our results show that P3 amplitude is positively linked to the match construct ($b=7.90$, $p<0.05$). Given the coding of the match construct based on Table 3 (i.e., 0 = mismatch and 1=match), P3 amplitude was lower in the match condition compared to the mismatch condition, and this confirms our expectations. To control for any learning effect, we also included the order in which the tasks were performed. Table 12 provides the results of the P3 amplitude regression model.

Table 12- Regression results for H1/ DV=P3 Amplitude (Mouse Click events)

P3 Amplitude (Mouse Clicks)	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	1.59	1.54	1.03	0.152	-1.446	4.627
Goal	-2.08	3.46	-0.60	0.273	-8.877	4.701
Sorting	-1.16	3.44	-0.34	0.367	-7.928	5.590
Match	7.90	3.45	2.29	0.011	1.141	14.663
Constant	-9.52	12.71	-0.75	0.227	-34.441	15.387
sigma_u	39.647					
sigma_e	15.031					
rho	0.874					

The second regression model with P3 latency as the dependent variable shows that a match between the user's goal and the product sorting variable significantly decreased P3 latency. As illustrated in Table 13, as the match variable increases (i.e., there is a fit between Goal and Sorting variables) the P3 latency decreases ($b=-33.99$, $p<0.01$); this is according to our expectations.

Table 13- Regression results for H1/ DV=P3 Latency (Mouse Click events)

P3 Latency (Mouse Clicks)	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	3.13	5.83	0.54	0.295	-8.288	14.564
Goal	-18.23	13.03	-1.40	0.081	-43.777	7.309
Sorting	-7.31	12.97	-0.56	0.286	-32.745	18.114
Match	-33.99	12.97	-2.62	0.004	-59.429	-8.555
Constant	474.71	34.02	13.95	0.001	408.03	541.398
sigma_u	28.021					
sigma_e	56.555					
rho	0.197					

We performed two other regression models similar to the first two ones but with P3 amplitude and latency of image presentation events as dependent variables. The third regression models shows that P3 amplitude is positively influenced by the match variable ($b=4.29$, $p<0.05$). P3 amplitude of the image presentation events is affected in a similar way to that of the mouse click events. As hypothesized, they both decrease as difficulty of the task increases. The regression results are presented in Table 14.

Table 14- Regression results for H1/ DV=P3 Amplitude (Image Presentation events)

P3 Amplitude (Image Presentation)	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	-1.55	0.86	-1.80	0.036	-3.245	0.140
Goal	0.52	1.93	0.27	0.392	-3.258	4.311
Sorting	0.31	1.92	0.16	0.434	-3.452	4.084
Match	4.29	1.92	2.23	0.012	0.526	8.065
Constant	30.07	5.22	5.76	0.001	19.834	40.309
sigma_u	7.256					
sigma_e	8.380					
rho	0.428					

The fourth regression model tests the effect of the match between the sorting variable and goal on the P3 latency of the image presentation events. As presented in Table 15, the P3 latency is negatively influenced by the match variable ($b=-40.31$, $P<0.01$), confirming our expectation that P3 latency increases with task difficulty.

Table 15- Regression results for H1/ DV=P3 Latency (Image Presentation events)

P3 Latency (Image Presentation)	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	5.04	5.94	0.85	0.198	-6.606	16.688
Goal	22.06	13.28	1.66	0.049	-3.976	48.098
Sorting	16.26	13.22	1.23	0.109	-9.658	42.184
Match	-40.31	13.23	-3.05	0.001	-66.242	-14.384
Constant	354.98	35.26	10.07	0.001	285.873	424.102
sigma_u	39.860					
sigma_e	57.649					
rho	0.323					

All four regression models show that P3 amplitude and latency are affected by the fit between sorting and goal variables. We further analyze the link between the match variable and cognitive load by testing the effect of the match variable on theta and alpha band powers. As stated, we expect that alpha power will decrease and theta power increase as cognitive load increases. Therefore, the next two regression models investigate how the match variable affects theta and alpha band powers.

The results presented in Table 16 show that alpha power is positively linked to the match variable ($b=0.73$, $p<0.01$), meaning that it is greater for match conditions compared to mismatch conditions. As we expected, alpha band power decreases as task difficulty increases.

Table 16- Regression results for H1/ DV= Alpha Band Power

Alpha Power	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	0.04	0.12	0.39	0.349	-0.181	0.271
Goal	0.30	0.26	1.15	0.125	-0.209	0.802
Sorting	-0.06	0.26	-0.23	0.409	-0.562	0.444
Match	0.73	0.26	2.85	0.002	0.229	1.236
Constant	4.13	0.89	4.65	0.001	2.387	5.872
sigma_u	2.590					
sigma_e	1.119					
rho	0.843					

The next model investigates the effect of the match variable on theta power.

Table 17 shows that Theta power is negatively influenced by the match variable (b=-0.83, P<0.001). It confirms our expectation that theta power increases as task difficulty increases.

Table 17- Regression results for H1/ DV= Theta Band Power

Theta Power	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	-0.17	0.11	-1.53	0.064	-0.388	0.048
Goal	-0.26	0.25	-1.04	0.151	-0.745	0.230
Sorting	0.15	0.25	0.59	0.279	-0.341	0.631
Match	-0.83	0.25	-3.36	0.001	-1.319	-0.347
Constant	5.58	0.76	7.33	0.000	4.087	7.073
sigma_u	1.926					
sigma_e	0.831					
rho	0.843					

All the six regression models show that user cognitive load, measured by P3 amplitude and latency of mouse click events, P3 amplitude and latency of image presentation events, and theta and alpha band powers, was less in match conditions compared to mismatch conditions. The two types of measures (i.e., ERP and Frequency analysis) reflect different aspects of the cognitive load construct. The ERPs are momentary measures of workload at the time of event presentation while the alpha and theta power yield an overall measure of cognitive load over a time window. Based on the evidence from both methods, H1 is strongly supported.

H2 suggests that cognitive load is negatively linked to task performance. Similar to the previous hypotheses, we performed six regression models to test the effect of cognitive load on performance. The first two regressions used P3 amplitude and latency of mouse click events as independent variables. The first model, tests the effect of P3 amplitude on task performance. Our results show that an increase in P3 amplitude (i.e. when cognitive load decreases) is associated with a decrease in the time required to perform a task, which means that performance increases ($b=-0.02$, $P<0.05$). The results are provided in Table 18.

Table 18- Regression results for H2/ IV=P3 Amplitude (Mouse Clicks)

Performance	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	0.24	0.24	0.990	0.162	-0.238	0.719
P3 Amplitude (Mouse Clicks)	-0.02	0.01	-1.820	0.035	-0.050	0.001
Constant	4.44	0.70	6.380	0.000	3.077	5.804
sigma_u	0					
sigma_e	2.419					
rho	0					

The second model uses P3 amplitude as the independent variable to test the effect of cognitive load on task performance. Our results show that as P3 latency increases (i.e., cognitive load increases), the time to perform a task increases as well, which means that performance decreases (b=0.01, P<0.01) as expected based on H2. The results are provided in Table 19.

Table 19- Regression results for H2/ IV=P3 Latency (Mouse Clicks)

Performance	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	0.16	0.24	0.68	0.250	-0.310	0.637
P3 Latency (Mouse Clicks)	0.01	0.00	2.17	0.015	0.001	0.017
Constant	0.27	1.87	0.14	0.443	-3.398	3.936
sigma_u	0					
sigma_e	2.355					
rho	0					

The next two models investigate the effect of P3 amplitude and latency of image presentation events on user performance. As provided in Table 20, the first model does not provide support for the expected relationship between P3 amplitude and user performance (b=-0.03, P =0.128).

Table 20- Regression results for H2/ IV=P3 Amplitude (Image Presentation)

Performance	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	0.16	0.25	0.62	0.266	-0.334	0.646
P3 Amplitude (Image Presentation)	-0.03	0.03	-1.14	0.128	-0.078	0.021
Constant	5.04	1.08	4.65	0.000	2.914	7.159
sigma_u	0					
sigma_e	2.407					
rho	0					

The next regression tests the link between P3 latency of image presentation events and user performance. Similar to the P3 latency of mouse click events, the results show that an increase in cognitive load (i.e. an increase in P3 latency) results in more time on the task and a reduction in performance (b=0.01, P<0.05). The results are provided in

Table 21.

Table 21- Regression results for H2/ IV=P3 Latency (Image Presentation)

Performance	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	0.17	0.24	0.71	0.239	-0.306	0.653
P3 Latency (Image Presentation)	0.01	0.00	1.67	0.048	-0.001	0.014
Constant	1.59	1.63	0.97	0.166	-1.615	4.785
sigma_u	0					
sigma_e	2.403					
rho	0					

Finally, we further investigate the relationship between cognitive load and user performance by using FFT measures of cognitive load. The next two models use theta and alpha band powers respectively as independent variables. As provided in Table 22, the first regression analysis shows that increasing theta power (i.e. increase in cognitive load) positively influences time on task (b=0.25, P<0.05).

Table 22- Regression results for H2/ IV= Theta Power

Performance	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	0.24	0.24	0.99	0.161	-0.234	0.713
Theta Power	0.25	0.13	1.84	0.033	-0.016	0.508
Constant	2.86	0.94	3.05	0.001	1.022	4.698
sigma_u	0					

sigma_e	2.132
rho	0

The last model shows that alpha power is negatively linked to user performance (b=-0.17, P<0.05). As cognitive load increases (i.e. alpha power decreases) the time required to perform the task increases as well, which signifies a reduction in user performance. The results for this regression model are provided in Table 23.

Table 23- Regression results for H2/ IV= Alpha Power

Performance	Coefficient	Std. Err.	Z	P Value (one-tailed)	95% conf. Interval	
Order	0.21	0.24	0.86	0.195	-0.265	0.678
Alpha Power	-0.17	0.10	-1.68	0.046	-0.378	0.029
Constant	4.93	0.84	5.87	0.000	3.286	6.577
sigma_u	0					
sigma_e	2.08					
rho	0					

Multiple evidence from different cognitive load measures shows that cognitive load negatively influences user performance in selecting the target product. Therefore, H2 is supported.

A summary of all the results for H1 and H2 is provided in Table 24.

Table 24- Summary of results for H1 and H2

H1 / Match ---------> Cognitive Load		
IV	DV	supported
Match	P3 Amplitude (Mouse Click Events)	✓
Match	P3 Latency (Mouse Click Events)	✓
Match	P3 Amplitude (Image Presentation Events)	✓
Match	P3 Latency (Image Presentation Events)	✓
Match	Alpha Power	✓
Match	Theta Power	✓
H2 / Cognitive Load ---------> Task Performance		
IV	DV	supported
P3 Amplitude (Mouse Click Events)	Task Performance	✓
P3 Latency (Mouse Click Events)	Task Performance	✓
P3 Amplitude (Image Presentation Events)	Task Performance	×
P3 Latency (Image Presentation Events)	Task Performance	✓
Alpha Power	Task Performance	✓
Theta Power	Task Performance	✓

2.5 Discussion

In this study, we argued that a match between a user's goal and product sorting on shopping websites would decrease the amount of cognitive load required to make a product decision. Results indeed show that users experience less cognitive load when the product sorting matches their goal (H1). This is in line with the general framework proposed by Todd & Benbasat (1992) on the effect of using computer-based decision aids on users' cognitive load. Sorting products is one instance of helping consumers to narrow down their consideration set and find their desired products. Consequently, they use less cognitive resources in making a product decision.

Two different EEG-based methods were used to measure cognitive load. The frequency analysis captures the total cognitive load that users experience during the selected time window (Mirhoseini et al., 2017). On the other hand, the ERP measure reflects the amount of cognitive load that the user experiences at the moment of event presentation. Although the two measures represent different types of cognitive load, we expected that they would be affected similarly by the task characteristics - specifically because the cognitive load is cumulative in nature and is affected by the task requirements of the past moments. Therefore, the ERPs should be influenced by the requirements of the task that was performed about 300 ms before the measurement point.

Our second hypothesis links cognitive load to task performance. Our results show that as cognitive load increases, task performance decreases, which is consistent with the framework proposed by Todd & Benbasat (1992) because cognitive load reflects how efficiently a task is performed by the users. P3 latency significantly predicts user performance. As P3 latency increases (i.e., cognitive load increases), user performance decreases (i.e., longer time spent on the task).

Our study contributes to research by showing the link between ERP components and user objective performance measures. Our results show that as the latency of the P3 component increases, user performance decreases (H2). It shows that as the load on working memory increases, it slows down the processing speed of the central executive unit indicated by the latency of P3, which consequently decreases user performance.

This research contributes to methodology by introducing a new method of measuring cognitive load during online shopping tasks. Similar to other ERP studies in the literature, which aimed at measuring cognitive load, we used the amplitude and latency of P3 components of the user's event-related potential. Two types of events were used to generate ERPs. The image presentation event is the moment that the task is presented to the participant. This type of event has been used in the past to measure cognitive load (Horst et al., 1984). However, unlike traditional ERP research, we also used one of the user's activities during the main task as an event to generate ERPs. In our experiment, we used the user's mouse clicks, which are natural events during the shopping session, to create ERPs. We believe that this design has potential for researchers in studying user-IT interactions for three reasons. First, mouse clicks are discrete events which can be segmented within a continuous user-IT interaction. They represent the user's cognitive response to the experimental task i.e. finding the target product. Second. they are temporally as close as possible to the user's decision moment. We can assume that the mouse clicks are the moments that users make their decision and locate the target product. Therefore, mouse click events show the user's cognitive status after making the decision. Finally, mouse clicks are natural events in an online shopping context. This type of design is more suitable for studying user behavior in a business context since it does not distract the user from the main task which is an interaction with information technology. Therefore, in studying user behavior in a user-IT interaction, the choice of events is crucial. Such events need to be discrete and temporally relevant to the phenomenon of interest.

This research also has implications for practice by showing how product sorting can reduce the user's cognitive load. Website designers need to offer users the type of sorting they desire in order to reduce their cognitive load and increase their performance. This is important since cognitive load is a predictor of user satisfaction in an online shopping context (lo Storto, 2013). It means that a no match condition reduces user satisfaction with the shopping experience.

Our research has limitations with respect to accounting for all possible user goals in online shopping. To be able to test our model, we limited our study to two simple goals. However, this opens avenues for future research to investigate other user goals. In this research, we studied how users with pre-defined goals (i.e., finding the cheapest product or finding a specific brand) are affected by a website's product sorting features. However, the constructive view of consumer decision making suggests that many users do not have a clearly predefined set of preferences to make product decisions (Payne, Bettman, Coupey, & Johnson, 1992). Their preferences and criteria for making a product decision are constructed in response to a number of tasks, contextual, and individual difference factors. Prior knowledge or expertise can affect the construction of individual preferences (Payne et al., 1992). Therefore, users who have no predefined strategy for decision making, may construct a set of preferences based on a number of factors. It is interesting to study how sorting products may affect users who do not have any predefined preference. Shopping websites also provide various type of filters for users to screen out the products that do not meet their minimum requirements. This feature is another instance of manipulating the presentation of information. If users engage in a shopping task without any predefined goal, how do different types of sorting or filtering affect their cognitive load and performance?

2.5 Conclusion

In this study, we investigate the contingent effect of product sorting and the user's goal on cognitive load. We argue that product sorting will decrease the user's' cognitive load in making a product decision if it matches the user's goal. Our study contributes to theory by uncovering the effect of the presentation order of information on the user's cognitive load. Another contribution is the demonstration of the mechanism by which information ordering affects task performance. Our analysis shows that providing the correct information ordering results in a more efficient use of working memory resources, which in turn increases task performance.

References

Aljukhadar, M., Senecal, S., & Daoust, C.-E. (2012). Using recommendation agents to cope with information overload. International Journal of Electronic Commerce, 17(2), 41–70.

Ariely, D. (2000). Controlling the information flow: Effects on consumers' decision making and preferences. Journal of Consumer Research, 27(2), 233–248.

Baddeley, A. (2007). Working memory, thought, and action (Vol. 45). OUP Oxford.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. Academy of Management Review, 3(3), 439–449.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. Journal of Consumer Research, 25(3), 187–217.

Cai, S., & Xu, Y. (2008). Designing product lists for e-commerce: The effects of sorting on consumer decision making. Intl. Journal of Human–Computer Interaction, 24(7), 700–721.

Carmel, E., Crawford, S., & Chen, H. (1992). Browsing in hypertext: A cognitive study. IEEE Transactions on Systems, Man, and Cybernetics, 22(5), 865–884.

DeStefano, D., & LeFevre, J.-A. (2007). Cognitive load in hypertext reading: A review. Computers in Human Behavior, 23(3), 1616–1641.

Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomics Science, 4(1–2), 113–131.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., & Rush, G. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. Human Factors, 40(1), 79–91.

Hale, S., Myerson, J., Rhee, S. H., Weiss, C. S., & Abrams, R. A. (1996). Selective interference with the maintenance of location information in working memory. Neuropsychology, 10(2), 228.

Häubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. Marketing Science, 19(1), 4–21.

Hong, W., Thong, J. Y., & Tam, K. Y. (2004). The effects of information format and shopping task on consumers' online shopping behavior: A cognitive fit perspective. Journal of Management Information Systems, 21(3), 149–184.

Horst, R. L., Munson, R. C., & Ruchkin, D. S. (1984). Event-related potential indices of workload in a single task paradigm. In Proceedings of the Human Factors Society Annual Meeting (Vol. 28, pp. 727–731). Sage Publications Sage CA: Los Angeles, CA.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual- task performance. Psychophysiology, 17(3), 259–273.

Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (1998). Extended ICA removes artifacts from electroencephalographic recordings. In Advances in neural information processing systems (pp. 894–900).

Käthner, I., Wriessnegger, S. C., Müller-Putz, G. R., Kübler, A., & Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. Biological Psychology, 102, 118–129.

Kleinmuntz, D. N., & Schkade, D. A. (1993). Information displays and decision processes. Psychological Science, 4(4), 221–227.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. Multiple-Task Performance, 279–328.

Léger, P.-M., Sénecal, S., Courtemanche, F., de Guinea, A. O., Titah, R., Fredette, M., & Labonte-LeMoyne, É. (2014). Precision is in the eye of the beholder: Application of eye fixation-related potentials to information systems research. Journal of the Association for Information Systems, 15(10), 651.

lo Storto, C. (2013). Evaluating ecommerce websites cognitive efficiency: An integrative framework based on data envelopment analysis. Applied Ergonomics, 44(6), 1004–1014.

Luck, S. J. (2012). Event-related potentials.

Mirhoseini, S. M. M., Léger, P.-M., & Sénécal, S. (2017). The Influence of Task Characteristics on Multiple Objective and Subjective Cognitive Load Measures. In Information Systems and Neuroscience (pp. 149–156). Springer.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology, 48(2), 229–240.

Müller-Putz, G. R., Riedl, R., & Wriessnegger, S. C. (2015). Electroencephalography (EEG) as a Research Tool in the Information Systems Discipline: Foundations, Measurement, and Applications. CAIS, 37, 46.

Murata, A., Uetake, A., & Takasawa, Y. (2005). Evaluation of mental fatigue using feature parameter extracted from event-related potential. International Journal of Industrial Ergonomics, 35(8), 761–770.

Oppenheim, A. V., Schafer, R. W., & Yuen, C. K. (1978). Digital signal processing. IEEE Transactions on Systems, Man, and Cybernetics, 8(2), 146–146.

Payne, J. W., Bettman, J. R., Coupey, E., & Johnson, E. J. (1992). A constructive process view of decision making: Multiple strategies in judgment and choice. Acta Psychologica, 80(1), 107–141.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(3), 534.

Postle, B. R., Idzikowski, C., Sala, S. D., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. Quarterly Journal of Experimental Psychology, 59(1), 100–120.

Schkade, D. A., & Kleinmuntz, D. N. (1994). Information displays and choice processes: Differential effects of organization, form, and sequence. Organizational Behavior and Human Decision Processes, 57(3), 319–337.

Sharkey, U., Acton, T., & Conboy, K. (2009). Modelling the effects of decision tools in online shopping. BLED 2009 Proceedings, 10.

Todd, P., & Benbasat, I. (1992). The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. Mis Quarterly, 373–393.

Uetake, A., & Murata, A. (2000). Assessment of mental fatigue during VDT task using event-related potential (P300). In Robot and Human Interactive Communication, 2000. RO-MAN 2000. Proceedings. 9th IEEE International Workshop on (pp. 235–240). IEEE.

Ullsperger, P., Freude, G., & Erdmann, U. (2001). Auditory probe sensitivity to mental workload changes–an event-related potential study. International Journal of Psychophysiology, 40(3), 201–209.

Ullsperger, P., Metz, A.-M., & Gille, H.-G. (1988). The P300 component of the event-related brain potential and mental effort. Ergonomics, 31(8), 1127–1137.

Ullsperger, P., Neumann, U., Gille, H. G., & Pietschann, M. (1986). P300 component of the ERP as an index of processing difficulty. Human Memory and Cognitive Capabilities, Amsterdam: Elsevier Science Publications.

Van der Heijden, H. (2006). Mobile decision support for in-store purchase decisions. Decision Support Systems, 42(2), 656–663.

Venkatraman, N. (1989). The concept of fit in strategy research: toward verbal and statistical correspondence. Academy of Management Review, 14(3), 423–444.

West, P. M., Ariely, D., Bellman, S., Bradlow, E., Huber, J., Johnson, E., … Schkade, D. (1999). Agents to the Rescue? Marketing Letters, 10(3), 285–300.

Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science, 3(2), 159–177.

Wickens, C. D., Isreal, J., & Donchin, E. (1977). The event related cortical potential as an index of task workload. In Proceedings of the Human Factors Society Annual Meeting (Vol. 21, pp. 282–286). SAGE Publications Sage CA: Los Angeles, CA.

Zeyl, T., Yin, E., Keightley, M., & Chau, T. (2016). Adding real-time Bayesian ranks to error-related potential scores improves error detection and auto-correction in a P300 speller. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24(1), 46–56.

Appendix1- EEG Data Analysis

- The MATLAB Code for filtering the EEG data:

```
FilterOrd=96;
CutOffFreq1=0.1;
CutOffFreq2=30;
SamplingFreq=500;
DataMatrix_Raw= EEGData;
bpFilt = designfilt('bandpassfir', 'FilterOrder', FilterOrd, 'CutoffFrequency1',
CutOffFreq1, 'CutoffFrequency2', CutOffFreq2, 'SampleRate', SamplingFreq);
DataMatrix_Filtered = filtfilt(bpFilt,DataMatrix_Raw);
```

This code was executed using the MATLAB transformation tool of Branvision Analyzer 2.0. This transformation first exports the data into MATLAB, executes the code and then imports it back to the Brainvision software.

Number of Segments per participant and the reason each segment was removed:

Table 25- EEG segments details

participant	Condition	Number of segments removed	Reason	Total Segments
1	1	0		10
	2	0		10
	3	0		10
	4	0		10
2	1	2	Artifacts	8
	2	0		10

	3	0		10
	4	0		10
3	1	0		10
	2	0		10
	3	0		10
	4	0		10
4	1	1	Wrong Answer	9
	2	0		10
	3	0		10
	4	0		10
5	1	0		10
	2	0		10
	3	0		10
	4	0		10
6	1	0		10
	2	0		10
	3	0		10
	4	0		10
7	1	0		10
	2	0		10
	3	0		10
	4	0		10
8	1	0		10
	2	0		10
	3	0		10
	4	0		10
9	1	0		10
	2	0		10

	3	0		10
	4	0		10
10	1	0		10
	2	0		10
	3	0		10
	4	0		10
11	1	0		10
	2	0		10
	3	0		10
	4	0		10
12	1	0		10
	2	0		10
	3	0		10
	4	0		10
13	1	0		10
	2	0		10
	3	0		10
	4	0		10
14	1	0		10
	2	0		10
	3	0		10
	4	0		10
15	1	0		10
	2	0		10
	3	0		10
	4	0		10
16	1	0		10
	2	0		10
	3	0		10

	4	0		10
17	1	0		10
	2	1	Artifacts	9
	3	1	Artifacts	9
	4	0		10
18	1	0		10
	2	0		10
	3	0		10
	4	0		10
19	1	1	Artifacts	9
	2	0		10
	3	0		10
	4	0		10
20	1	0		10
	2	0		10
	3	0		10
	4	0		10

Chapter 4-
Thesis Conclusion

Conclusion

Based on two essays, this thesis investigated the measurement and use of the cognitive load construct in information systems design research. We argued that complications in measuring cognitive load have been preventing IS researchers from using it to evaluate and improve the design of IT artifacts. This thesis is expected to contribute to IS research by introducing different metrics of cognitive load that can be used in various design studies.

NeuroIS is a subfield in information systems which aims at informing IS researchers using neuroscience tools and theories (Loos et al., 2010). In this thesis, one of these tools (EEG) was used to propose new metrics of cognitive load. Thanks to the high temporal resolution of EEG, we were able to measure the user's instantaneous workload in an online shopping experience. When we extracted three metrics from it, accumulated load, peak load, and average load, our results showed that accumulated load is the most comprehensive measure among them since it was the only measure that could capture the effect of both factors that contribute to workload (Task difficulty and task uncertainty). As cognitive load is associated with different behavioral constructs such as user satisfaction and emotion (Gwizdka, 2010), it can be used to evaluate the design of IT artifacts. High temporal resolution of workload metrics such as accumulated load and peak load allows us to study the efficiency of IT artifacts in a natural way, and avoid traditional biases associated with the use of subjective measures.

A second experiment was designed as a continuation of the first experiment to test the convenience of an online shopping interface. We hypothesized the effect of search convenience on accumulated load. Convenience reflects user efficiency in online shopping. An inconvenient shopping website will have two consequences: increase in shopping time and increase in mental effort. The accumulated load construct captures both of these factors (Mirhoseini et al., 2016), therefore, it can be used to assess user convenience in any stage of the shopping process. We also hypothesized the effect of accumulated load on user satisfaction. We expect that this study will contribute to IS research by establishing accumulated load as a criterion for designing more convenient IT artifacts.

In the second essay, we studied the contingent effect of product sorting and user goal on the user's cognitive load. Information display is an important aspect of online shopping websites because it affects users behavior (Schkade & Kleinmuntz, 1994). We argued that if the user's goal (the specific criteria that the user has for choosing a product) is aligned with product sorting, it helps the user make product decisions more easily and save working memory resources. To capture a user's cognitive load in we used two types of cognitive load measures: 1- Event-Related-Potentials (ERP) and 2-Frequency analysis. We also used different events in order to provide multiple evidence for the effect of the fit between the user's goal and product sorting on the user's cognitive load. We further investigated the effect of cognitive load on user performance. Therefore, we expect that our third essay contribute to IS research by explaining the contingent effect of a user's goal and product sorting on user mental workload. It also contributes to research by using multiple measures of cognitive load in an online shopping context and proposing a new way of evaluating IT design elements.

References

Gwizdka, J. (2010). Distribution of cognitive load in web search. Journal of the American Society for Information Science and Technology, 61(11), 2167–2187.

Léger, P.-M., Sénecal, S., Courtemanche, F., de Guinea, A. O., Titah, R., Fredette, M., & Labonte-LeMoyne, É. (2014). Precision is in the eye of the beholder: Application of eye fixation-related potentials to information systems research. Journal of the Association for Information Systems, 15(10), 651.

Loos, P., Riedl, R., Müller-Putz, G. R., vom Brocke, J., Davis, F. D., Banker, R. D., & Léger, P.-M. (2010). NeuroIS: neuroscientific approaches in the investigation and development of information systems. Business & Information Systems Engineering, 2(6), 395–401.

Mirhoseini, S., Léger, P.-M., & Senecal, S. (2016). The influence of task characteristics on multiple objective and subjective cognitive load measures. Presented at the Gmunden Retreat on NeuroIS.

Schkade, D. A., & Kleinmuntz, D. N. (1994). Information displays and choice processes: Differential effects of organization, form, and sequence. Organizational Behavior and Human Decision Processes, 57(3), 319–33