Sampo Noponen

ARE MEMORIES CONSISTENT WITH EXPERIENCE? EXPLORING TEMPORAL ASPECTS OF HEDONIC WEBSITE UX

JYVÄSKYLÄN YLIOPISTO
TIETOJENKÄSITTELYTIETEIDEN LAITOS
2019

ABSTRACT

Noponen, Sampo
Are memories consistent with experience? Exploring temporal aspects of hedonic website UX
Jyväskylä: University of Jyväskylä, 2019, 104 p.
Information Systems, Master's thesis
Supervisors: Kyppö, Jorma; Silvennoinen, Johanna & Kujala, Tuomo

This study explores the progression of single user experience over a short-term period in the context of websites. Depending on temporal distance between user and event of interaction, user experience can be divided into three phases: before, during and after interaction. Momentary emotions of pleasure and pain experienced during interaction largely determine the memory of user experience. Memories, however, do not always correspond with actual experience, and consequently people can form judgements and decisions that do not reflect their experience. Such memory biases might originate from sequencing effects: certain moments of an episode are more impactful to formation of retrospective evaluations than others. Due to the above reasons, this thesis examined how consistent hedonic retrospective website user experience evaluations tends to be with real-time evaluations. Thesis investigated influence of certain sequencing effects, such as peak-end effect, on retrospective evaluations. In recent years, practitioners have been increasingly interested in utilising animated transitions into interfaces. Yet there is scant research addressing how they influence user preferences. Therefore, this study additionally explored influence of animated transitions on user experience. Results were acquired with empirical user experiments where participants interacted with a website and quantitatively evaluated their experience. As it turned out, retrospective evaluations were mainly consistent with real-time evaluations, despite minor differences. Thus, it can be inferred that memory was consistent with actual experience. Strong peak-end, and recency effects were found, although effect size decreased in time. Lastly, animated transitions did not influence user experience. Explanations for these results are considered and future research possibilities are proposed.

Keywords: user experience, hedonic evaluation, website, design, sequencing effects, animated transition

TIIVISTELMÄ

Noponen, Sampo
Ovatko muistot yhdenmukaisia suhteessa kokemukseen? Tarkastelussa verkkosivun hedonisen käyttäjäkokemuksen ajallinen näkökulma
Jyväskylä: Jyväskylän yliopisto, 2019, 104 s.
Tietojärjestelmätiede, Pro gradu -tutkielma
Ohjaajat: Kyppö, Jorma; Silvennoinen, Johanna & Kujala, Tuomo

Tässä työssä tarkastellaan yksittäisen verkkosivun käyttäjäkokemuksen kehityskulkua lyhyellä aikavälillä sekä käyttöliittymän siirtymien vaikutusta kokemukseen. Riippuen käyttäjän ja vuorovaikutustilanteen ajallisesta etäisyydestä toisiinsa, yksittäinen käyttäjäkokemus voidaan jakaa kolmeen vaiheeseen: 1.) ennen vuorovaikutusta, 2.) vuorovaikutuksen aikana sekä 3.) vuorovaikutuksen jälkeiseen vaiheeseen, jolloin jäljellä on vain muisto. Kokemisen aikana koetut tunteet määrittävät pitkälti sen minkälainen muisto kokemuksesta jää. Muistot eivät kuitenkaan ole aina yhdenmukaisia sen kanssa mitä todella tapahtui. Tästä seurauksena ihmiset voivat tehdä kokemuksista epäjohdonmukaisia arvioita tai päätöksiä. Aiemmissa tutkimuksissa on huomattu, että tällaiset muiston vääristymät voivat johtua sekvensointiefekteistä: toiset hetket vaikuttavat takautuviin arviointeihin muita enemmän. Näistä syistä johtuen, tässä tutkielmassa tarkasteltiin kuinka yhdenmukaisia käyttäjien takautuvat verkkosivun käyttäjäkokemuksen hedoniset arvioinnit ovat suhteessa reaaliaikaisiin arviointeihin. Tutkittiin, kuinka paljon sekvensointiefektit, kuten peak-end efekti, vaikuttavat takautuviin arviointeihin. Tutkimuksessa selvitettiin lisäksi animoitujen siirtymien vaikutusta käyttäjäkokemukseen. Viime vuosina animoituja siirtymiä on enemmissä määrin käytetty käyttöliittymissä pehmentämään käyttöliittymän tilojen muutoksia. Tästä huolimatta animoitujen siirtymien vaikutusta käyttäjän mieltymyksiin ei ole juurikaan tutkittu. Tutkimus suoritettiin empiirisesti käyttäjäkokeilla, joissa osallistujat käyttivät verkkosivua sekä arvioivat kokemuksiaan. Tulokset osoittivat, että huolimatta pienistä eroista, takautuvat kokonaisarvioinnit olivat pääasiassa yhdenmukaisia reaaliaikaisten arviointien kanssa. Muistot täten noudattivat kokemusta. Lisäksi kokemuksen huippukohta ja loppukohta korreloivat vahvasti takautuvien arviointien kanssa. Kuitenkaan animoitujen siirtymien vaikutus käyttäjäkokemukseen ei ollut merkittävä. Tuloksien syitä analysoidaan, ja aiheita jatkotutkimukselle tarjotaan.

Asiasanat: käyttäjäkokemus, hedoninen arviointi, verkkosivusto, suunnittelu, sekvensointiefektit, animoidut siirtymät

FIGURES

TABLES

CONTENTS

TIIVISTELMÄ
ABSTRACT
KUVIOT
TAULUKOT

1 INTRODUCTION

User experience (UX) is a complex phenomenon. In fact, it is so complex, it might be impossible to take account of every single factor affecting it (Jokinen, 2015, p. 16). Many designers are ignorant of this complexity, and hence prone to make erroneous decisions that decrease perceived quality of designed products (Norman, 2010). Designers should not blindly trust their intuition. A characteristic of human cognitive system is that it oversimplifies judgements about the world, and even experts might overconfidently trust their judgements in cases where they actually do not have reason for such confidence (Kahneman, 2011, p. 209). This illusion of validity can be prevalent for UX practitioners as well, who produce solutions in their complex profession.

Nevertheless, the better picture we have of user psychology and nature of subjective experience, the better equipped we are designing enjoyable experiences (Hassenzahl, 2010, p. 29). As there are, perhaps, too many factors to consider for one designer, prioritisation is necessary. Most importantly, he or she should know central aspects of experience, ones that crucially determine whether experience was perceived as positive or not. This paper adopts Hassenzahl's (2008) view on central aspects of UX. According to this view, fulfilment of basic needs has a decisive role in positive experiences.

Why should practitioners go to the trouble of learning psychology just for inducing more pleasure to human-computer interactions (HCI)? Because that can improve our quality of life. Like its cousin paradigm, positive psychology, UX has a novel goal: to increase well-being of the world (Hassenzahl & Tractinsky, 2006). Life consists of series of experiences. Hence, goal of improving well-being can be attained by providing enjoyable experiences. This master thesis attempts to take a step towards that goal by examining how individual website experiences are formed temporally.

Speaking of temporal UX, critical distinction can be made between *experiencing* and *an experience* (Hassenzahl, 2010, p. 19; Forlizzi & Battarbee, 2004). Experiencing refers to actual moment when interaction is happening, while an experience refers to state when interaction has ended and only thing that remains from it is the memory. Distinction is important, as a designer should primarily design for the memory of the experience rather than for the actuality,

the time span when interaction is happening (Cockburn, Quinn, & Gutwin 2017; Karapanos, Zimmerman, Forlizzi & Martens, 2010; Norman, 2009). Memory of an experience is what matters when user decides whether to use the service again or not (Cockburn et al., 2017).

Memory contains far less information than what actually happened during interaction (Kahneman, 2011, p. 381). As an implication, if we consider UX only as a 'memory of experience', the concept shrinks to something less complex. However, designing for memories might be even more challenging than designing for actuality. For instance, perceptions during experiencing cannot be affected by events that occur after interaction, whereas memory of experience can be affected (Anderson, 2015, p. 154). Usually, events that occur during interaction primarily determine memory of experience (Hassenzahl, 2008). Yet, one should not neglect potential influence of events occurring before and after interaction on memory.

> They may forget what you said, but they will never forget how you made them feel. (Carl W. Buehner, 1971)

Above mentioned quote from Carl W. Buehner contains key to experiences: salience of emotions. Quote maintains that when everything else is forgotten, last thing left from a memory are the emotions one felt. This is not merely a folk-psychological observation anymore. Scientists found that patients suffering from Alzheimer's disease felt prolonged elevated levels of sadness even though they could not remember a single factual detail regarding the experience (Guzmán-Vélez, Feinstein & Tranel, 2014). Moreover, intensity of emotions influences level of detail, vividness and confidence in memory (Tambini, Rimmele, Phelps, & Davachi, 2017). Clearly, memories and emotions are closely interrelated (Plutchik & Kellerman, 2013, p. 8).

Salience of emotions holds true for *user* experiences as well (Hassenzahl, 2010, p. 3). Momentary emotions experienced during interaction chiefly determine the memory of UX (Hassenzahl, 2008). Imagine a person who has once bought an e-book from certain e-commerce site and is pondering whether or not to visit that site again. Undoubtedly, she will make the decision based on a feeling associated with the memory of her first experience with the site (Hassenzahl, 2008). If the feeling is positive, she just might buy e-books again in future. If it is negative, she will not visit that site any time soon. People tend to avoid negative to a greater extent than they embrace positive (Kahneman, 2011, p. 282).

As memories can be biased, would it thus be possible that objectively positive experiences can be remembered as negative or vice versa? Several psychological experiments have extensively proved that this can indeed be the case (e.g., Kahneman, Fredrickson, Schreiber and Redelmeier, 1993; Redelmeier & Kahneman, 1996b). Against one's intuition, experiments demonstrated that average of momentary feelings experienced during an episode does not often correspond particularly well with retrospective summary evaluation. Kahneman et al. (1997) have attributed this phenomenon to *peak-end effect*, which refers to overweighed influence of the most intensive moment and final moment of epi-

sode to memory. This view is supported by the fact that experiences that elicit arousal are more likely to be remembered than experiences that do not (Kensinger, 2009).

Memory bias can be measured by comparing real-time evaluations (instant utility) of an experience to retrospective evaluations (remembered utility) of the same experience (Kahneman et al, 1997). Corresponding to experiencing, *real-time evaluation* refers to assessment of valence and intensity of current hedonic experience (Kahneman et al, 1997; Kahneman, 2000). Real-time evaluation cannot include considerations of whole episode, but only one moment of it. Corresponding to an experience, *retrospective evaluation* is defined in present thesis as overall assessment of a past experience (Kahneman et al., 1997). It refers to a process of retrieving memory of experience for assessing how good it was. It is important to understand how retrospective evaluations are formed in human cognitive-affective system. This knowledge is highly valuable to many practitioners, who would benefit from increased ability of designing positive UX memories. But ultimately this knowledge would benefit users, whose well-being would be positively influenced by improved UX designs.

On account of aforementioned reasons, dynamic temporal dimension of UX is a subject ripe for scientific inquiry. In present thesis, objective is to find out how consistent are users' real-time evaluations compared to their retrospective evaluations of a single UX in a laptop website context. We are interested in how momentary feelings felt during experiencing translates to retrospective summary assessments. Moreover, aim is to study whether sequencing effects, especially peak-end effect, influence users' memories of experiences, and retrospective evaluations. Does the order of moments pertaining to experienced episode cause possible inconsistencies? Expectedly, as so much of the past experiences is lost, biases are inclined to occur. Importance of peak-end effects is well noted in the field of psychology, but only few such studies has been conducted in HCI research (Cockburn et al., 2017).

Layout transition is another temporal element of UX. Animated transitions, which constitute one particular type of user interface (UI) animation, are nowadays commonly used in mobile applications and operating systems. Animated transition is a smooth change between two visual states of interface (Chevalier et al., 2014). Such a transition can be observed by unlocking a smartphone. Locked view does not instantly change to unlocked view. Instead, smoothly yet swiftly enough, a new view emerges from the bottom of the screen, and old view slides upwards until it has completely disappeared. By opening any major application, one is likely to witness instances of animated transition. While they are utilised to a lesser extent in the web, animated transitions are increasingly common among modern websites. Considering the prominent role they play in modern interfaces, there exists surprisingly little empirical research on how animated transitions actually influence users' perceptions of interaction (Merz et al., 2016).

Animations and transitive elements of interface can be considered to belong to both pragmatic and hedonistic side of experience (Hassenzahl, 2004; Merz et al., 2016). Researchers have theorized that animated transitions can bring about ease of understanding (e.g., Chang & Ungar, 1995; Thomas & Cal-

der, 2001), user engagement (Chevalier et al., 2014) and aesthetic pleasure (Merz et al., 2016; Huhtala et al., 2010). They are potential source of playfulness and pleasure, by enriching the interaction with novel and surprising elements. Furthermore, animation can propel cognitive benefits, as Bederson and Boltman (1999) found out that animated transitions help users to build mental maps of spatial information. Properly designed, transitions can guide users' attention (Merz et al., 2016). Importantly, only appropriate use of animation can improve experience (Merz et al., 2016). Especially in work-related interfaces, adding 'unnecessary' elements to interfaces is likely to only negatively impact experience (Maeda, 2016).

Theorizations aside, research has not been able to demonstrate how animated transitions influence real-time nor retrospective evaluations of UX thus far. Hence, that is another area of interest in this paper. More specifically, aim is to examine whether their influence is positive, negative or non-existent. Also, it is interesting to investigate relationship between animated transitions and sequencing effects.

Thesis is conducted as an empirical study where quantitative data is gathered from user experiments. During experiment, participant is asked to perform several tasks with a goal-oriented and non-work-related laptop website. Study is conducted with mixed study design, including both within-subjects and between-subjects study. Participants are divided into two groups. Experimental group uses website with animated transitions, whereas control group uses the same website without said transitions. In both groups, participants evaluate their website experience twice. First measurement point takes place during experiencing and right after experience. That is when momentary hedonic evaluations are measured. Second measurement point takes place one week after the interaction when retrospective evaluations are measured. Momentary evaluations correspond with how participants actually felt during experience, while retrospective evaluations reflect how well participants *remember* how they felt during experience.

1.1 Objectives and research questions

Purpose of this study is to gain new knowledge on temporal aspects of UX, how consistently users evaluate experiences, and how animated transitions influence these evaluations in the website context. Present study has potential to reveal how perceptions of UX experiences evolve in short-term period. There is a gap in understanding how events happening during experiencing lead to formation of affective UX. It is uncertain, which moments of experience are most critical in formation of subsequent memories of experiences. These objectives will be accomplished by answering the following research questions:

1. How consistent are hedonic retrospective UX website evaluations to real-time evaluations*?*

2. Can retrospective evaluations be predicted with peak-end rule (and other sequencing effects)?
3. How animated transitions influence UX right after interaction and one week after interaction?

First question is the main research question of the thesis, whereas second and third serve as additional research questions. Present thesis adopts Hassenzahl's (2008) user-centred approach and Kahneman's et al (1997) experienced utility theory for evaluating UX of websites. Results of this study should be beneficial for UX designers in understanding phenomenon of UX more profoundly and guide them to most important aspects of UX.

1.2 Outline

In here, contents of the paper are briefly covered. The work begins with introduction where background, goals, topics, and motivations for the study are presented. Introduction offers an overview of the thesis for a reader. This study includes theoretical and empirical parts. Chapter 2 describes how theoretical part of the study was conducted by the means of literature review. Afterwards, results of literature review are presented in chapters 3 and 4. Important concepts are explained in detail and prior research on study topics is presented. Chapter 3 is devoted to analysis of phenomenon of UX. It begins by defining concept of UX and explaining its main aspects. Thesis will adopt Hassenzahl's (2008) view on UX. Additionally, more attention is given to temporal aspects. UX is divided into phases and sequencing effects are explored. Literature review continues in fourth chapter, where animations and animated transitions are scrutinized, and prior research on animated transitions is presented. Emphasis is put on animated transitions' role in website context.

In chapter 5, methods used in empirical part of the study are presented. Empirical research model, derived from the findings of literature review, is described firstly. It is followed by descriptions of participants, stimuli, variables, procedure of data collection and data analysis. In chapter 6, results of the study are presented elaborately. Chapter 7 may be the most important part of the thesis. It contains discussion where findings of the study are compared with prior research. Furthermore, implications and contributions of present thesis to UX theory, UX practice and UX research methodology are considered. Lastly, conclusions, limitations and possibilities for future research are described in chapter 8.

2 METHOD OF LITERATURE REVIEW

Results of literature review are presented in third and fourth chapter of this paper. This chapter, however, contains description of how literature review was conducted. It can be roughly divided into two parts, which are review of prior UX research and review on animated transitions. The former part demanded more time, as it contained more information to absorb. Literature review had multiple objectives. First one was to acquire general overview of the nature of UX. General view was necessary precedent for understanding more specific aspects of experience, and important for providing context for the present study. Aim was to get to the bottom of what UX is. As a by-product of general review process, suitable UX model was found (Hassenzahl, 2008), and it was adopted to serve in the empirical part of the study. Secondly, temporal quality of UX had to be analysed as a background for research topic of memory biases and sequencing effects. Third objective was to review and synthesize prior HCI research on sequencing effects and animated transitions for finding a research gap. Fourth objective was to find how consistency between instant utility and remembered utilities should be measured in website experience context. Fifth objective was to get acquainted with phenomenon of animated transitions and review prior findings of HCI studies on animate transitions. Lastly, prior literature was scrutinised in the hope of finding advices on how animations should be implemented. Ultimately, literature review provided research model for the empirical part of this paper.

Procedure of literature review began with keyword search. Keyword search method stands for "querying of quality scholarly databases by the use of a specific word or phrase" (Levy & Ellis, 2006). Source material was searched and found solely by using internet service Google Scholar. Primarily keyword queries were conducted without using advanced search function. Intention was to accept articles from the most credible publishers, but due to lack of research both on sequencing effects and animated transitions, researcher had to lower threshold at times. Above all, articles were accepted based on their relevance to the topic. The second-most important selection criterion was amount of citations per article, or in other words, acceptance among researchers. Other criteria were article title, article abstract, and article's order among search results and

apparent quality of its contents. Although keyword search method yielded many articles, most of the publications referred in this paper were found with method of backward references search. Backward references search refers to process of reviewing references of the articles found with keyword search (Levy & Ellis, 2006). Source material is mainly from HCI research, but also from fields of psychology, economics, information systems science, neuroscience, cognitive science and information visualization.

Nowadays there is abundance of UX literature. Therefore, only major HCI articles, books or authors were accepted as source material for presenting overview on UX in chapter 2.1. Firstly, articles were searched with keyword query 'user experience'. It yielded enormously 5 300 000 results. First results were Hassenzahl and Tractinsky's (2006) article, Garrett's (2010) book, and study of Law et al. (2009). They provided a good starting point for the review. Present thesis adopted Hassenzahl's (2008) model as it fits well with Kahneman's (1997) experienced utility theory. Hassenzahl's (2008) was sixth result of the query 'user experience'. Hassenzahl et al. (2010) and Karapanos et al. (2009) were ninth and tenth. Most of the articles were found with forward or backward searching references of aforementioned articles. Forward references search is a process of reviewing additional articles that have cited the article (Levy & Ellis, 2006). Law et al. (2009) provided most of the definitions listed in table (TABLE 1). Forlizzi and Battarbee (2004) was found with backward searching Hassenzahl's (2008) article. Articles regarding temporal aspects of UX were searched with query 'UX over time'. Articles regarding user evaluation measurement were searched with query "measuring affective state".

Literature review on prior HCI research on sequencing effects was conducted by using Cockburn's et al. (2017) article. The article was found with keywords '"sequencing effects" and "HCI"' which yielded 29 results. Article was the first option among the results. All the other HCI articles pertaining to this topic were found with backward search of Cockburn et al. (2017). This article was highly suitable because it is very recent, from the major HCI publisher (International Journal of Human-Computer Studies), and its topic could not be much more relevant for present study. It was used as an anchor for finding all the relevant studies to this specific research area of sequencing effects.

Psychological research articles on peak-end effects were mainly found with backward search of Cockburn et al. (2017), which summarizes prior research on sequencing effects in both psychology and HCI research. The main article by Kahneman et al. (1997) was found with backward search of Cockburn et al. (2017). Additionally, forward and backward authors search was used. These methods refer to reviewing what authors have published before or after the article (Levy & Ellis, 2006). One of the leading authors of UX is Mark Hassenzahl, and Daniel Kahneman is leading scholar in psychological research on peak-end effects. Kahneman's book (2011) was simply found with query "Kahneman" from Google Scholar, it being the first result among 214 000 options.

Sources for animated transitions were searched with queries *"website" AND "animated transition"* (395 results), *"user experience" AND "animated transition"* (310 results), *"HCI" AND "animated transition"* (146 results), and *animated*

transition (217 00 results). Most of the results did not seem relevant for the context of website UX or up-to-date. Websites have advanced drastically from the beginning of millennium, so it would be important to find current articles. Merz et al. (2016) article, which was ninth result with query *"user experience" AND "animated transition"*, was chosen as starting point. Though it is but a late-breaking article with few citations, it was chosen because its title was especially relevant for this thesis, included major scholars from UX field, and is fairly new. Its references provided good amount of source articles for the present study. For instance, backward reference search revealed major papers of Chang and Ungar (1995), Thomas and Calder (2001), and Thomas Johnston (1981). Chevalier et al (2014) was first result with query *"HCI" AND "animated transition"*.

When most relevant articles regarding sequencing effects and animated transitions were found, they were analysed with the following process:

1. Articles were read thoroughly from the abstract to the conclusions, one by one.
2. Especially relevant lines and sentences were highlighted.
3. When reading process was finished, short one-page summary of the article was written.
4. Summaries were categorized based on their topic.

Written summaries contained article's topic, objective, short description of chosen method, summation of results and, additionally, a critique of article. Process of writing summaries was time-consuming, but it provided a solid overview on research topics. After all summaries had been finished, author felt time was appropriate for writing down results of literature to chapters 2 and 3. Statements brought forward in chapters were supported with supplementary articles. Vast majority of the source articles were of supplementary nature and were analysed with more straightforward process. Supplementary articles were skimmed through, without reading the entire contents of articles.

3 USER EXPERIENCE

Here, the key concept of this thesis, namely UX, is defined and analysed. While many aspects of UX are ignored, focus is put to aspects deemed as most relevant for the present work. We are especially interested in emotions, memory and temporal dynamics of UX. Present paper adopts Hassenzahl's (2008) user-centred approach to UX and UX design. Firstly, general overview to the UX is offered. Secondly, Hassenzahl's (2008) model is presented. Thirdly, readers are introduced to temporal perspective to UX, including sequencing effects. Lastly, study briefly takes a look on how UX has been measured in HCI studies, and which presented method could be most suitable for the needs of present study.

3.1 What is user experience?

Some time ago, when a practitioner needed an answer to tricky question of how one succeeds in acquiring satisfied and long-term users of their products, a researcher answered simply: make the product as efficient, effective and satisfactory as possible and test the product to ensure its performance (Hornbæk, 2006). If someone were to ask the same question now, the answer will be more complicated: do not just focus on the product, pay attention to users and context as well (Hassenzahl & Tractinsky, 2006); understand user needs and design interface in a manner that it satisfies those needs (e.g., Norman, 2013, p. 6; Hassenzahl, 2008); understand users' motivations (Hassenzahl, 2008); comprehend importance of emotions (e.g., Jokinen, 2015, p. 14; Hassenzahl, 2010, p. 3; Norman, 2004, p. 5); ensure that interface provides aesthetic pleasure (e.g., Tractinsky, Katz & Ikar 2000; Lavie & Tractinsky, 2004; Moshagen & Thielsch, 2010); acknowledge that UX changes over time (Karapanos et al., 2009); make a good first impression (Lindgaard, Fernandes, Dudek & Brown, 2006); provide fun and playful interactions (Hassenzahl, 2010, p. 41); do not design too complex solutions, but not too simple neither (e.g., Maeda, 2006; Rousi & Silvennoinen, 2018); and do not forget: make that interface usable (Hassenzahl, 2008; Norman, 2013, p. 5).

Purpose of previous paragraph is to illuminate differences between two major paradigms of HCI research, namely usability and UX. Usability can be regarded as a quality attribute of UI, which regards certain pragmatic aspects of interaction, such as efficiency and frustration minimisation. As a counter-movement to usability, UX tries to incorporate everything that the interaction experience entails, including instrumental pragmatic aspects and non-instrumental hedonic aspects (Hassenzahl & Tractinsky, 2006.). While usability is very important, it is but a one part of UX.

Relatively recent trend in psychology research helps us to comprehend another great difference between two grand paradigms of HCI. Traditionally psychology research has restricted itself to mental problems, to the extent that Seligman and Csikszentmihalyi (2000, p. 5) claimed that "psychologists have scant knowledge of what makes life worth living". Claim might be unjust exaggeration, but it highlights the problem — positive aspects of and well-being has been largely neglected. And thus, positive psychology paradigm was born with a mission to change focus of psychology from preventing pathologies to subjective and communal well-being (Seligman & Csikszentmihalyi, 2000, p. 5). Instead of battling mental suffering, emphasis is put to positive aspects of life, such as happiness and well-being (Hassenzahl, 2010, p. 28.). Acknowledging Kahneman and Tversky's prospect theory (1977), it is not surprising that researchers have been inclined to avoid negative instead of approaching positive. Rule of loss aversion dictates that humans prioritize avoiding losses over seeking gain — even if it would be more beneficial to do the opposite (Kahneman & Tversky, 2013).

UX in HCI research is counterpart of positive psychology in psychology research (Hassenzahl, 2010, p. 28). Before, interaction research focused on preventing occurrence of negative outcomes, emotions and experiences. UX, however, takes a different approach to interaction. Alike positive psychology, it focuses on positive aspects of life, such as fun, joy and pleasure, and adopts them to HCI context. Yet UX does not forget pragmatic aspects of interaction. Daringly, UX paradigm aims for fulfilment through well-designed technology (Hassenzahl, 2010, p. 29).

UX is a holistic approach to HCI (e.g., Hassenzahl, 2010 p. 11; Forlizzi & Battarbee, 2004). That is to say, it acknowledges all elements pertaining to the formation of experience of interaction. Hence, it is no wonder that concept of UX is hard to pin down. Scholars have given many different definitions to this buzzword of recent years, of which several have been listed in the table below (TABLE 1). Most of the listed definitions have academic origins, but there are also two definitions originated from industry (Norman & Nielsen, 2018; ISO, 2008). Some definitions emphasise certain aspects of experience, while other definitions focus on other aspects. None of them succeed in depicting holistic overall picture of UX. Actually, UX seems to be too multifaceted entity to be captured with a single theory (Jokinen, Silvennoinen & Kujala, 2018). However, by considering multiple definitions together, we are able to form more comprehensive picture of UX.

TABLE 1 Definitions of user experience

Author(s)	Definition
Hassenzahl & Tractinsky (2006)	" A consequence of a user's internal state (predispositions, expectations, needs, motivation, mood, etc.) the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organizational/social setting, meaningfulness of the activity, voluntariness of use, etc.)".
ISO (2008)	"A person's perceptions and responses that result from the use or anticipated use of a product, system or service".
Norman & Nielsen (2018)	"User experience encompasses all aspects of the end-user's interaction with the company, its services, and its products."
Desmet & Hekkert (2007)	"The entire set of affects that is elicited by the interaction between a user and a product including the degree to which all our senses are gratified (aesthetic experience) the meanings we attach to the product (experience of meaning) and the feelings and emotions that are elicited (emotional experience)."
Sward & Macarthur (2007)	"The value derived from interaction(s) [or anticipated interaction(s)] with a product or service and the supporting cast in the context of use (e.g. time, location, and user disposition)."
Hassenzahl (2008)	a. "A momentary, primarily evaluative feeling (good-bad) while interacting with a product or service" b. "Good UX is the consequence of fulfilling the human needs for autonomy, competency, stimulation (self-oriented), relatedness, and popularity (others-oriented) through interacting with the product or service (i.e., hedonic quality). Pragmatic quality facilitates the potential fulfilment of be-goals."

Noticeably, that there are many ways to define UX, depending on how one approaches the topic. Law, Roto, Hassenzahl, Vermeeren and Kort (2009), have noticed that some definitions focus on company that provides UX (e.g., Norman & Nielsen, 2018), some on user which interacts with the product (e.g., Desmet & Hekkert, 2007), and others focus on evaluation of experiences (e.g., Hassenzahl & Tractinsky, 2006). Different definitions focus on different temporal dimensions, for instance, during or after interaction (Law et al., 2009.).

Scopes vary as well. While Hassenzahl (2008a) simplifies and narrows UX down to feelings that interaction inflicts in user, Norman and Nielsen (2018) embrace the broadness and complexity of UX. Furthermore, latter definition implies that UX is not only limited to single product, it also to encompasses whole interplay between customer and company. They are not the only ones who concur to this notion (Law et al., 2009). However, this supposition is closer to brand experience, which is associated to customer's general view on a company (Law, 2009). For the sake of clarity, in present thesis UX is confined to products, services and systems, which contain a user interface (Law et al., 2009).

Albeit distinct from each other, each definition entails insightful notions. For instance, Norman and Nielsen's (2018) definition illustrates that direct interaction does not only shape the experience, but also anything that is related to interaction between user and product can have an influence. Author's experi-

ence with iPod serves as an anecdotal example of this. In 2008 Apple had succeeded to market their iPod Nano music player extremely successfully to the point that author felt urgent need to buy the device. Their product was desirable to the extent that, having unboxed the device, author experienced inappropriately elevated feelings even before he had turned device on. Furthermore, these elevated feelings certainly increased perceived enjoyment when author finally got to interact with the device.

Lallemand, Gronier & Koenig (2015) surveyed 758 UX professionals on their viewpoints regarding UX, and found that among five definitions, UX practitioners ranked Hassenzahl & Tractinsky's (2006) definition as preferred option. Hence, it is commonly accepted that UX is a combination of three contributing agents: a user, a system and a context (Hassenzahl & Tractinsky, 2006). First agent, a user, entails psychological aspects of interaction. While there is no UX without an interactive product, it is incorrect to assume that UX is characteristic of a product. Instead, UX is psychological phenomenon that occurs in user's cognitive and affective processes informed by characteristics of the product (Silvennoinen, 2017, Jokinen et al., 2018). Experience is a complex phenomenon that arises from the integration these processes, such as action, perception, motivation and emotion — all of which happen inside user's head (Hassenzahl, 2010, p. 16). Therefore, UX is a subjective phenomenon. Even though two people would use same product in the same context at the same time, they could perceive the experience very differently from each other. Each person perceives the world in their own way.

There will be no UX without the second agent, that is the interactive system, or object. Having a walk in the woods is an experience, but it is not a *user* experience. This distinction is important to make in HCI. UX arises only when user interacts with some object. UX researchers have used various study objects, such as websites, mobile apps, work-related applications, or other forms of information systems (Law et al., 2009). However, interactive object can be many things: a restaurant, a washing machine, an art object or a door. Even the process of writing present paper can involve experience design. What are characteristics of potential readers? How should one structure the content so that reader comprehends it with ease? Which terms and word choices suit best in this particular academic context? Each object has its own characteristics that shapes experiences. Is the object designed for work-context or for entertainment? What functions it has? In the next chapter, relationship between user characteristics and characteristics of designed system is delved further. To be clear, in present thesis we are mainly interested in website as an interactive artifact.

Last contributing agent, the context, is often overlooked, yet important. It refers to the environment, or the situation, where interaction happens. Factors such as culture, time and location can greatly alter the nature of experience. If interaction happens in a workplace user might have different expectations and be in different mood than if the interaction would happen in home or café. In short, context influences internal state of user (Hassenzahl & Tractinsky, 2006).

Forlizzi and Battarbee (2004) have identified three types of experience: experiencing, an experience and co-experience. *Experiencing* is "constant stream of self-talk that happens when we are conscious" (Forlizzi & Battarbee, 2004). It

emphasizes the ever-present nature of experience — it is there whether we like it or not (McCarthy & Wright, 2004, p. 50).

Certain episodes from this constant stream are more significant, meaningful and emotionally evocative than others. They stand out from the rest of the stream as memories. Each of these standout episodes is *an experience* (Forlizzi & Battarbee, 2004). An experience has a beginning and end, and it can invoke emotional and behavioural change (Forlizzi & Battarbee, 2004). It is stored to experiencer's long-term memory and integrated to her knowledge of the world (Hassenzahl, 2010, p. 17). While constant stream of experiencing is ever-present, an experience is in the past. It is a completed and labelled episode (Forlizzi & Battarbee, 2004). In HCI studies, researchers are more interested in an experience than experiencing per se (Hassenzahl, 2010, p. 5).

In *co-experience,* more than one person is involved, and meaning and emotions are crafted together through the use of product (Battarbee & Koskinen, 2005). It refers to social aspects of experience. At times people form experiences mediated with products alone, however, quite often people create experiences with others. A subjective solitary experience can be co-experience as well if it is retrospectively shared with others (Forlizzi & Battarbee, 2004). Humans are intrinsically social animals (Tomasello, 2014), and thus they tend put more significance to shared experiences over solitary experiences. As a research topic, co-experience has been overlooked in HCI (Battarbee & Koskinen, 2005). State of affairs is not different in present study, where emphasis is put to experiencing and an experience, while co-experience is largely disregarded.

UX can be measured with user evaluations (Bargas-Avila & Hornbæk, 2011). In this thesis, we view evaluation as a subjective and qualitative assessment directed to certain experience. Furthermore, we presume that user evaluations are mainly derived from the affective appraisal mechanism of human mind (Hassenzahl, 2008). Simply put, users evaluate experiences based on how they feel about them. As Hassenzahl (2008) states, these feelings serve as 'psychological currency' that allows comparison of different experiences. A single evaluation contains two dimensions of emotion: valence (or pleasantness) and intensity (or arousal). Only these two dimensions have been reliably found from empirical studies of emotional experience (Smith & Ellsworth, 1985). Real-time evaluation refers to self-report assessment, which is given during experiencing. However, retrospective evaluation is an assessment which is given by the user only after the experience has ended. In order to assess the experience retrospectively, user has to retrieve that certain episode from memory. UX could be measured with more objective measures, such as psychophysiological instruments, which measure user's bodily reactions to stimuli. However, we chose self-reports over psychophysiological measures since, as of now, former method is deemed as the most suitable way of gaining insight of emotional experience (Jokinen, 2015, p. 53).

Different scholars take different approaches to UX. For instance, there are user-centred, product-centred and interaction-centred models made to understand experience (Forlizzi & Battarbee, 2004). Straightforwardly, *product-centred* approach focuses on providing information that enables creating high quality products that evoke enjoyable experiences. *User-centred* approach helps practi-

tioners to understand users. User-centred models integrate knowledge from different disciplines to understand experiential and psychological aspects that are especially relevant to UX. *Interaction-centred* approach emphasises the role products play in bridging the gap between designers and users. (Forlizzi & Battarbee, 2004.).

Alike UX definitions, none of the UX approaches fully captures the phenomenon of UX (Jokinen et al., 2018). Nevertheless, they are helpful for both business and academic purposes. This paper will adopt Hassenzahl's (2008) user-centred model of UX, since, while UI plays an important part in formation of UX, in present thesis we are more interested in user psychology. As we shall see, Hassenzahl's (2008) view regarding emotions fits well with our data collection methods. This model, which links product attributes to human needs and emotions is further elaborated in the following sub-chapter.

To conclude, UX is a sum of every aspect that has part to play in the interplay between human and technology. It is a combination of three agents: a user, a system and a context. It is subjective phenomenon that arises from the integration of cognitive-affective processes. As we do not yet have complete understanding of how cognitive-affective system in our mind functions, consequently it is impossible to have a complete comprehension of the nature of UX. It is both an episode that can be measured in time and a memory of that episode that is stored in person's mind. There are three types of experience, experiencing, an experience and co-experience. UX can be understood by taking different approaches, such as product-centred, user-centred and interaction-centred approaches. UX paradigm has a goal of bringing about more pleasure and happiness to world through good experiential design. It focuses on positivity.

3.2 User-centred model of UX

Before we lay out Hassenzahl's model (2008), we need to acquire some understanding of emotions. Emotion is another complex concept whose definition lacks consensus among academics (Jokinen, 2015, p. 46). Alike other functions pertaining to living organisms, emotions seem to have originated through evolutionary adaptive mechanism of natural selection, as put forth by Darwin and Wallace (1858), for the purpose of promoting survival of species (e.g., Izard, 1992; Bradley, Codispoti, Cuthbert, & Lang, 2001). Emotions serve many different vital functions for everyday life such as social communication, mobilisation for action and attention (Bradley et al., 2001). Commonly people regard emotions as opposite for rationality. This notion is partly misleading, since, as they are biochemical mechanisms for increasing likelihood of survival, emotions can be seen as embodiments of evolutionary rationality (Plutchik & Kellerman, 2013, p. XV). Surely, making decisions based on emotional responses can at times lead to undesirable results, as in business context for instance. But this does not imply that emotions are essentially irrational, rather mechanism of emotions was not developed for the modern business context.

Ekman and Cordaro (2011) have defined emotions as "discrete, automatic responses to universally shared, culture-specific and individual-specific events". Emotions contain five main components, which are cognitive appraisal processes, neurophysiological processes, action tendencies (motivational component), motor expression, and subjective feeling (Scherer, 2009). Characteristics of emotions are rapid onset, short duration, unsolicited occurrence and automatic appraisal (Ekman, 1992b) Emotions have biological origins, but emotional system can change greatly over the course of individual's life (Plutchik & Kellerman, 2013, p. XV).

Basic emotion theory states that evolutionary process has given birth to set of basic emotions that differ from each other in tone, such as happiness, sadness, fear, and surprise (Ekman, 1992b, Saarimäki et al., 2015). They differ from each other in manner of expression, behavioural response, and physiology (Ekman, 1992b). However, academics have not found consensus regarding number of these emotions. For instance, Ekman (1992a) states that there are at least five different basic emotions, while Smith and Lazarus (1993) maintain that there are four core emotions. Saarimäki et al. (2015) have identified six basic emotions that differ from each other neurally and physiologically. Each basic emotion is not a single affective state, but a family of relative states. Confusion regarding the number of basic emotions can be partly attributed to failure to distinguish emotion families from emotion variations within family (Ekman, 1992b.).

Regardless the exact number, it is useful to simply categorise emotions to pleasurable and painful (Plutchik & Kellerman, 2013, p. 7). Emotions are signals of survival, which can be described with approach-avoidance principle. The principle states that approach behaviour is towards something of positive survival value whereas avoidance behaviour is away from something of negative survival value (Plutchik & Kellerman, 2013, p. 5.). In similar vein, appraisal theory of emotion postulates that subjective emotional experience stems from cognitive process of appraisal (Silvennoinen, 2017, p. 36). Appraisal is a process of evaluating personal significance of an event (Jokinen, 2015, p. 55). Hence, a pleasant emotion emerges from subjective evaluations of an event that is deemed to have positive survival value for the person, and vice versa unpleasant emotion emerges from event that is evaluated as having negative survival value. As shall be shortly seen, appraisal theory fits well with Hassenzahl's (2008), and Hassenzahl and Tractinsky's (2006) theories of UX.

In this paper, at one place there has been talk about emotions and elsewhere about feelings. To clarify, that this is not done for arbitrary reasons. *Emotions* and *feelings* should not be considered as synonymous with each other. Details between these terms differ among theories, but commonly emotions are considered to represent the whole affective processing, while feelings are emotions that are consciously acknowledged (Jokinen, 2015, p. 50). Emotions can occur unconsciously, but when they are consciously perceived, they qualify as feelings (Prinz, 2005). One of the main disagreements among scholars pertain to problem of dualism: the relationship between conscious experience of emotion and physiological bodily basis of emotion (Jokinen, 2015, p. 46).

Be that as it may, emotions constitute one of the most central aspect to UX (Hassenzahl, 2010, p. 3). There are vast range of activities that may or may not

occur during experience, but one element is always present, which is momentary feeling of pleasure or pain in various intensities (Hassenzahl, 2008). According to appraisal theory of emotion, momentary feeling is an emotional response that depends on three aspects: a person who encounters it, perceived stimulus, and circumstances of the encounter (Jokinen, Silvennoinen & Kujala, 2018). If we compare this to Hassenzahl and Tractinsky's (2006) definition of UX, we see how well these two theories are aligned: UX is a consequence of user's internal state (a person who encounters), characteristics of the system (perceived stimulus) and the context (circumstances of the encounter).

In the end, that emotional response is what counts when a person thinks should she end or continue current episode. As discussed above, feelings of pleasure and pain can be derived from approach-avoidance principle. When something causes unpleasant feelings in us, our instinct tells us to retreat. If something makes us feel good, we want more of it. If the feeling is pleasant, one is likely to continue episode. If it is unpleasant enough, the opposite action takes place. Hence, emotions perpetuate action. Usually, positive experience demands positive emotions (Hassenzahl, Diefenbach & Göritz, 2010). The question is, how exactly positive emotions emerge during interaction? Or rather, what aspects of interaction determine the judgement of appraisal process? For answering the question, one needs to turn to actions, goals and human needs.

In short, Hassenzahl's (2008) model asserts that good experience is a consequence of fulfilled needs. Need fulfilment leads to arousal of positive feelings, which in turn leads to positive experiences. Need fulfilment has positive survival value for the person. On the other hand, when a person wants something to happen, but it does not happen, pain and negative feelings emerge. Similar to emotions, aspiration to fulfil needs is motivational trigger for action (Hassenzahl et al., 2010 p. 39). That is, actions are being ignited by the motivation to reach some goal. Motivation can be intrinsic or extrinsic (Hassenzahl, 2010, p. 39). When product use is initiated by external factor, UX is not likely to be perceived as positive as in a situation when product use originates from user's own desires (Hassenzahl, 2010, p. 39). Thus, it is important for designers to know where the motivation to use their products originates from. Moreover, sometimes people do actions because there is value in the process, for example jogging. Other actions are done as value is in the product, such as in DIY projects (Garrett, 2010.).

Interacting with technology can be seen as goal-directed action (Hassenzahl, 2010, p. 11). Carver and Scheier (1989) suggest that goals have three fundamental levels that control activities. A holistic UX consists of these three levels of goals: be-goals, do-goals and motor-goals (Hassenzahl, 2010, p. 13). An interactive product should enable user to accomplish all these goal for producing positive experience.

On the highest level of the goal hierarchy, there are *be-goals*. They are the most abstract level of goals. Be-goals are self-referential: they dictate motivations, purposes, and emotions behind experiences (Hassenzahl, 2008.). They relate to fundamental questions such as: What one would like to be? How one would like to feel about oneself? These goals give meaning to interaction. Thus, they are the most important level of hierarchy. They can be seen as synonymous

to basic human needs. For instance, a teenager posts a selfie photograph to a social media platform, such as Instagram, because he has be-goal of feeling himself important and special. He also feels need to achieve be-goal of being related to others. In second, and more pragmatic level, there are *do-goals*. They refer to concrete results of interaction. Getting a selfie uploaded to Instagram is a do-goal. While achievement of do-goal presents instrumental means to accomplish be-goal, *motor-goal* presents instrumental means to accomplish do-goal. Process of uploading selfie to a social media can be divided to different sub-goals, that is motor-goals. In this case, a single motor-goal is searching and picking that particular image file from the phone memory that is desired to be sent to app's server. In order, to accomplish this goal, teenager has to scroll down through images and tap the finger at the right image.

According to Hassenzahl's (2008) model, UX can be divided to pragmatic and hedonic qualities. *Pragmatic quality* refers to object's perceived ability to assist user in accomplishing aforementioned do-goals and motor-goals of interaction. It focuses on the object of the interaction, and its usability and utility. Consistently to usability paradigm, HCI researchers, and supposedly numerous IT practitioners, have been mainly focused on do-goal achievement (Hassenzahl, 2008). This is slightly concerning as before turning to pragmatic aspects, focus should be targeted to *hedonic quality* of interaction. A usable, well-functioning product serve no one if it does not fulfil any basic needs of users. Hedonic quality focuses on the user itself. It refers to product's perceived ability to accomplish be-goals. Before using a product, there has to be reason why it is used. Likely that reason is derived from basic human needs. Hedonic qualities relate to perceived non-instrumental value that interface generates to user, for example, by evoking feelings of fun and aesthetic pleasure. Mekler and Hornbæk (2016) take a step further by distinguishing hedonic experience to refer momentary pleasure derived from the interaction, whereas eudaimonic experience refers to accomplishment of personal goals and need fulfilment related to personal growth. In the experiment of present thesis, we will be principally collecting evaluations of hedonic instead of eudaimonic experience.

There exist several distinct psychological needs with varying degree of importance to UX. Hassenzahl et al. (2010) found that among universal psychological needs, there are seven that are especially relevant to UX, namely competence, relatedness, popularity, stimulation, meaning, security and autonomy. These seven needs, accompanied by descriptions taken from Sheldon, Elliot, Kim and Kasser (2001), are listed in the table (TABLE 2). Some of these needs, relatedness, popularity and relatedness, were already addressed in selfie-example. Hassenzahl et al. (2010) go even further and claim that experiences can be categorized by the primary need they aim to fulfil. It could be beneficial for UX designer to begin project by identifying experience category for the designed product. Is the product aimed for fulfilment of simulation, popularity, or perhaps for competence? Characteristics of designed product will depend on this initial identification.

TABLE 2 Psychological needs relevant to UX

Need	Description
Autonomy–Independence	Feeling like you are the cause of your own actions rather than feeling that external forces or pressure are the cause of your action
Competence–effectance	Feeling that you are very capable and effective in your actions rather than feeling incompetent or ineffective
Relatedness–belongingness	Feeling that you have regular intimate contact with people who care about you rather than feeling lonely and uncared of
Influence–popularity	Feeling that you are liked, respected, and have influence over others rather than feeling like a person whose advice or opinion nobody is interested in
Pleasure–stimulation	Feeling that you get plenty of enjoyment and pleasure rather than feeling bored and under stimulated by life
Security–control	Feeling safe and in control of your life rather than feeling uncertain and threatened by your circumstances
Self-actualizing–meaning	Feeling that you are developing your best potentials and making life meaningful rather than feeling stagnant and that life does not have much meaning

Results from studies are consistent with Hassenzahl's (2008) model. Hassenzahl et al. (2010) studied positive user experiences with 548 participants and found clear relation between need fulfilment and positive emotions. Especially needs stimulation (.44), relatedness (.26), competence (.29) and popularity (.24) were highly correlated with positive emotions. In another study, where researchers studied over thousand reports of positive experience, results support the finding that more intense the need fulfilment is, more positive the overall experience is (Hassenzahl, Wiklund-Engblom, Bengs, Hägglund & Diefenbach, 2015). Furthermore, need fulfilment was related to hedonic but not to pragmatic quality perceptions. In other study, researchers wanted to get to the bottom of why people return interactive products (den Ouden, Yuan, Sonnemans & Brombacher, 2006). Results showed that half (48 %) of the returned products were fully functional. Why would anyone return functional products? As it turned out, in 28 % of the cases, products were returned because they failed to satisfy user needs. Partala and Kallinen (2011) studied satisfying and unsatisfying experiences and found that needs of competence and autonomy were most salient in positive experiences and missing in negative experiences. Highly importantly, study results suggest that most satisfying user experiences are related to personally meaningful aspects of UX, whereas the most unsatisfying experiences were related to direct emotional responses of pragmatic problems. Their study concurred with theory that positive emotions lead to positive experiences.

Aforementioned findings do not mean that great experiences cannot or should not include negative emotions as well. Sometimes negative emotions are even needed for the formation of rich and positive experiences (Hassenzahl, 2010, p. 29; Silvennoinen, 2017 p. 39). Consider, for example, a film Schindler's list. It depicts a story of business owner who gets concerned over his Jewish

workforce during World War II. Among other things, the film proceeds to reveal truly gruesome horrors of concentration camp. Because of its highly distressing scenes, an average viewer cannot go through the movie without experiencing negative emotions. Yet it is one of the most loved movies of all time (IMDB). Undoubtedly, screenwriters of the film are not the only ones in their field that are aware of the fact that good dramatic experiences necessitate tension originating from the interplay of positive and negative emotions.

The task of experience designer is to create products that have instrumental value. In other words, products have to fulfil user's needs. Depending on situation, urge to fulfil needs vary. At some point one does not have any need to feel competent, but to feel related to others. Context, time and personal characteristics influence which need of the list is more relevant than others. More prolonged the fulfilment of need is, more likely there is an urge to fulfil that need. The urge arises from need deprivation (Hassenzahl, 2010, p. 38). Without pragmatic qualities of the product, such as functionality and usability, psychological need of the user cannot be fulfilled. In other way, without hedonic qualities of the product, functionality and usability serve no purpose (Hassenzahl, 2010, p. 13.). Thus, a designer has to combine both qualities to craft an excellent product.

To summarise, Hassenzahl's model (2008) describes one perspective to UX. It lays out a theory of how practitioners can succeed in evoking positive and compelling experiences. It views emotions, actions, and human needs as central parts of UX. Positive UX is consequence of arousal of positive emotions. Positive emotions are aroused when interaction successfully satisfies one or more of the seven psychological needs that are relevant to UX. Urge to fulfil these needs, such as stimulation, relatedness, and competence is a motivational trigger to the action of interacting with a product. In order to provide satisfactory experiences, designer has to ensure both hedonic and pragmatic quality of a product. Product of high pragmatic quality does not have any purpose if one cannot fulfil any need with it. Product of high hedonic quality falls short of providing good UX, if its usability is terrible.

3.3 Temporal perspective on UX

Like every phenomenon in this universe, UX is not a static, for it changes over time (e.g., Hassenzahl, 2010; Karapanos, Zimmerman, Forlizzi & Martens, 2009; Kujala, Vogel, Pohlmeyer & Obrist, 2013). Passing of time can alter completely one' perception of the product. For instance, Karapanos et al. (2009) studied iPhone owners, and found that pleasure of novelty decreased sharply after few weeks of use. Time is closely connected to with memory, as over time our memories of past experiences can change and even biases can occur.

Many HCI scholars acknowledge the importance of time in UX (e.g., Hassenzahl, 2008; Forlizzi & Battarbee, 2004; Karapanos et al., 2009). Additionally, UX practitioners agree that UX changes constantly when users interact with product (Lallemand et al., 2015). Respondents in the same study also agreed

that momentary real-time evaluations are favoured over retrospective evaluations of UX. Contrary to this notion, in present chapter, arguments are given to assert the view that retrospective evaluations are more significant and helpful for the needs of UX practitioners. When experiencing, present moment is everything one has, but as all experiences end sooner or later, what really matters is the memory of interaction.

Regarding contents of present chapter, we begin by showing how UX can be divided to different temporal phases. Afterwards, prior psychology research on peak-end effect is reviewed. Kahneman's et al. (1997) utility theory is adopted to explain this particular bias of memory. When acquainted with psychological background, prior HCI research on sequencing effects is examined. Lastly, relationship between two types of experiences, experiencing and an experience, is analysed and their importance is compared.

3.3.1 Three phases of a single user experience

Length, frequency and amount of individual experiences with a product can vary greatly. In some cases, people interact with products somewhat frequently during a long-term period. This happens when a person needs to use online banking system to manage her finances. In some cases, people interact with products extremely frequently over a long period of time. For example, normally people interact with their mobile phone dozens of times every day, and this user-product relationship can last few years. Sometimes, user only interacts with a product only once in their lifetime. This likely occurs when a product provides a negative experience. From the business perspective, it is important for many companies to convert people to regular and frequent users of their products, as their income depends on it. From aggregate of individual experiences overall broad UX is formed. However, in present thesis we are interested in a single interactive episode between user and product, as even one interaction can be crucial to the future success of a product. Before we know how perceptions of UX evolve during time, one needs to understand the role of a single experience.

Depending on temporal distance between a user and an event of interaction, a single UX can be divided broadly to three phases: before, during and after interaction (Law et al., 2009; Roto, 2007). First phase starts when user becomes aware of forthcoming experience, but the actual interaction is still in the future. It is related to phases of expected UX (Roto, 2007) and anticipation (Karapanos et al., 2009). Before interaction, expectations towards forthcoming experience are formed, and these expectations can alter the nature of experience. They are formed from the information that user connects to the product. Expectations can develop consciously or unconsciously. Negative words from a friend will lower expectations while positive stories raise expectations. Raita and Oulasvirta (2010) tested how expectations (of positive or negative valence) influenced post-experiment usability ratings and found, unsurprisingly, that positive expectations positively influence perceived usability. Interestingly enough, this effect remained even when users failed all the tasks during exper-

iment. Karapanos et al. (2009) found that in long term, actual experiences can influence more to perceived user satisfaction than their expectations before interaction.

Before interaction, expectations are not the only element, that can alter forthcoming episode. Recent events happened prior interaction set user's mental state (Hassenzahl & Tractinsky, 2006). These events might not have anything to do with the interaction, but they can greatly affect how the user perceives the interaction. For example, if a person hears that her friend is stricken with serious disease before interaction, it is unlikely that person will be able to enjoy or even pay much attention to interaction during experiencing. Unfortunately, often designer's abilities to control user's mental state are exceedingly limited.

Undoubtedly, second temporal phase, UX during interaction, is utmost important to formation of overall UX. That is the phase where the actual phenomenal events pertaining to single interactive episode occur. Simultaneously, user subjectively perceives occurring events through sense organs and interprets them through cognition (Hassenzahl, 2010, p. 19). During interaction user is involved with experiencing. From the designer's point of view, it is vital to know key factors of interaction that determine how recollected experience is perceived. The whole episode of interaction can be further divided to separate temporary parts.

First impressions are created during the first moments, or better, milliseconds of the interaction. First immediate moments during the interaction have potential to greatly influence the overall perceived experience (Lindgaard, Fernandes, Dudek & Brown, 2006). We can consider first impressions as something that occurs in a limbo: an episode has already started but user has barely interacted with the product yet. First impressions regard largely aesthetics of interaction. Studies on first impressions have mainly investigated aesthetics, but there is some indication that users can rapidly form reliable judgements of other product characteristics as well, such as trust and credibility (Tuch, Presslaber, StöCklin, Opwis & Bargas-Avila, 2012). Lindgaard et al. (2006) studied how quickly participants form their opinion of perceived aesthetic appeal of a web page. They found that users formed reliably first impressions in less than 500 milliseconds (half a second). However, Thielsch and Hirschfield (2012) demonstrate that due to methodological problems, Lindgaard et al. (2006) likely overestimated the effect of first impressions. Additionally, Thielsch and Hirschfield (2012) show that ultra-rapid aesthetic judgements are only connected with low-spatial frequencies. It means that fine details are not relevant for ultra-rapid judgements, underlying overall layout is. In any case, first impressions are of high importance, as they can determine whether user will continue interacting with the product.

Interaction can be divided to sub-interactions, or moments, each of which yield different outcomes (Cockburn et al., 2017). A useful way to divide interaction is based on the most meaningful moments of interaction. Meaningful sub-interactions can be found from the studies of sequencing effects, such as primacy, recency and peak-end effects. Sequencing effects refer to notion that during experiencing, the order of sub-interactions affect person's memory of an experience (Cockburn et al., 2017). According to sequencing effects, the most mean-

ingful parts are initial, ending and the most intense moments of interaction. First impressions belong to initial moments of interaction. Sequencing effects are covered thoroughly in the chapters 2.3.2 and 2.3.3.

After interaction, being third and last temporal phase of single experience, starts after episode of interaction has ended and a person is left with the memory of experience. Memory is dynamic, that is, information associated with it susceptible for alteration over time (Hassenzahl, 2010, p. 20). One week after experience, memory of UX might be different than what it was immediately after experience. According to decay theory of forgetting, in typical experience, the more time has passed since it happened the more difficult it gets to retrieve details of the experience (Anderson, 2015, p. 152–154). This relationship between memory retention and time is called the power law of forgetting (Anderson, 2015, p. 153). When an experience is particularly special to a person, he or she is likely to think about it more often, which leads to better memory retrieval and more details are available for retention. This can be explained with the power law of learning. Simply put, the power law states that when a memory is practiced, memory activation is strengthened (Anderson, 2015, p. 138).

Memory of experience can also alter due to interference from other memories (Anderson, 2015, p. 154). As surmised above, other people's opinions heard before interaction can change how the interaction will be experienced. But it is also possible that opinions heard *after* interaction can affect perceptions of experience. Learning additional associations can cause old associations to be forgotten (Anderson, 2015, p. 155), and as a result, memory will change. For instance, after playing through a videogame someone else can mention a fact that presents the game in a new light. This new fact about the game is remembered while some old fact is forgotten. Consequently, that fact has changed player's view of the experience.

To conclude, a single experience can be broadly divided to three temporal phases, namely before, during and after interaction. Events occurring in any of these phases can alter the perceived quality of the experience. Initial mental and emotional state can influence how one will perceive the forthcoming episode. Likewise, advent of new information can influence how one views past episode. Nevertheless, in most cases it is the events occurred during experience that most crucially affect the memory of the experience.

3.3.2 Sequencing effects in psychology research

Intuition tells that memories of experiences are fairly consistent with the actual experiences. But are they actually? At this point it is worth clarifying that in this study we are not interested in how accurately people remember factual details of memories. Instead, emphasis is placed upon how well people remember feelings that were felt during experience. In the field of psychology, from where the study topic has originated from, there has been fair amount of research regarding sequencing effects, especially peak-end effect. In numerous occasions, it has been proved that peak-end effect can cause inconsistencies between momentary and retrospective evaluations of an experience (e.g., Ariely, 2008; Do, Rupert & Wolford, 2008; Kahneman et al., 1993; Redelmeier & Kahneman, 1996a).

Psychologists have applied term experienced utility for assessing the perceived value of experiences. The term is derived from the Bentham's concept of utility (1789), which holds that pleasure and pain are two 'sovereign masters' that fundamentally determine human actions. Kahneman, Wakker & Sarin (1997) have formed a formal theory for experienced utility of temporally extended outcomes (TEOs). TEO's can be seen synonymous to what user experiences are in HCI research. A TEO is an experience.

Kahneman et al.'s (1997) theory distinguishes different forms of utility that concern experiences. *Decision utility* is "a measure of an experience which is inferred from choices" (Kahneman et al., 1997). It is a weight of an outcome of decision. If one has two options from which to choose from, she will choose option with higher perceived decision utility. Decision utility does not include any considerations of hedonic states. The term has been adapted to decision theory in economics. Aforementioned *experienced utility* is a hedonic quality of an experience, which is associated with outcome of an experience (Kahneman et al., 1997.). It is an emotional response that reflects the global evaluations of an episode on a pleasure-pain scale. Single evaluation entails the valence (good or bad) and the intensity (from mild to extreme) of affective or hedonic experience. If experience is deemed as immensely pleasurable and joyful, such as memory of a wedding day for many, the experience has very high experienced utility. In contrast, traumatic experiences have high experienced disutility. While it is a hedonic quality of experience, pragmatic qualities included in episode can influence experienced utility.

Kahneman et al. (1997) argue that experienced utility can be measured with subjects' retrospective evaluations of total pleasure or displeasure, which is called as *remembered utility*, or with momentary reports of current subjective experience, namely *instant utility*. As stated in chapter 2.2., pleasure and pain reflect adaptive functions that are biologically programmed to our brains. Remembered utility corresponds to retrospective evaluations of an experience. It measures what a person remembers from their experience of past event (Cockburn et al., 2017). Remembered utility has an adaptive function of determining whether past experience should be avoided or approached in the future (Kahneman et al., 1997). While remembered utility is directed to past episode, instant utility corresponds to momentary real-time evaluations. Aggregated reports of instant utility construct *total utility*. Total utility represents objective overall quality of an experience. Contrary to remembered utility, order in which sequences are experienced does not affect total utility (Kahneman, 2000). Remembered utilities can be measured with three ways: (1) "reported evaluations of past experiences, (2.) physiological indications of emotion aroused by reminders of event", or (3.) with approach or avoidance tendencies they induce. Lastly, authors distinguish *predicted utility,* which refers to beliefs about experienced utility of outcomes.

Term 'utility' is a bit troublesome, as it has been used to mean many different things. As a matter of fact, Bentham himself apologised for using this word in context of experiences (Kahneman, 2011, p. 377). He just could not find a better word for the concept. In HCI research, one might connect utility with usability and task-oriented measures, instead of evaluations of overall experi-

ence. Moreover, utility is not very commonly used term in UX research. Due to these reasons, in present thesis terms real-time and retrospective evaluation are used for assessing the quality of UX, instead of terms total utility and remembered utility.

Results of studies on remembered utility support the notion of peak-end effect (Kahneman et al., 1997). When evaluating experienced utility, human cognitive system does not calculate objectively total utility of an experience, instead it takes short cuts. *Peak-end effect* is a sequencing effect, where user's memory is strongly influenced by the most intense moment of the interaction and the end of the interaction (Kahneman, 2011, p. 385). Along the same line, *peak-end rule* predicts that retrospective evaluations of an experience are formed as average level reported of the best (or the worst) moment of experience and its end (Kahneman, 2011, p. 380). Peak-end rule can be considered as psychological heuristic (Gutwin, Rooke, Cockburn, Mandryk & Lafreniere, 2016), and a straightforward way to measure experiences.

Peak-end effect occurs because our memory has evolved to represent most intense moment of episode and its end (Kahneman, 2011, p. 385). Humans do not remember past life consistently and in a linear fashion (Harrison, Amento, Kuznetsov & Bell, 2007). When we are experiencing, we process large quantities of information. Through our senses we register actions and events that occur around us in our environment. With our ears we hear noises, voices, music and talk. With our eyes we read text, watch videos, recognize people and navigate through traffic. With our hands we touch objects. With our nose we smell different odours from close proximity. Human brain is not capable of capturing and storing every instance of every single moment to its memory, which is why there has to be some logic for prioritising events. From the evolutionary perspective, it seems the most beneficial strategy is to prioritise emotional moments. Psychological studies have confirmed that emotionally intense events are more likely remembered than mundane events (Kensinger, 2009). Emotional arousal during an experience enhances subsequent retention of that experience, resulting in vivid detailed and rich memories (Tambini et al., 2017). Key player in this phenomenon appears to be amygdala, an almond-shaped cluster of nuclei, where emotional information is processed. Amygdala enhances retention of emotionally intensive memories by modulating hippocampus in storing arousing events (Phelps & Sharot, 2008.). Interestingly enough, being vivid, detailed and rich does not mean that emotional memories are necessarily *accurate*. Instead, people often remember details of highly emotional memories incorrectly (Phelps & Sharot, 2008).

Let us imagine a person had an experience of buying movie tickets online. A trigger for the experience was that she and her husband had prolonged urge to fulfil need of stimulation. It was Friday, and they wanted to feel themselves entertained after a long and dull week of work. Before interacting with the website, she was feeling relaxed, and with the exception of very bad end, interaction with a movie site went well and smoothly. She was easily able to find the wanted movie theatre amongst all theatres that company provide, the right movie amongst all movies available, and the correct showtime amongst all options. As a pleasant surprise, website's graphical seating chart enabled her to choose

great seats from the centre of the theatre. Moreover, interface seemed modern and even visually appealing. However, after filling all the billing information and when she proceeded to pay the tickets, web service got stuck in limbo of loading screen until it crashed. Person's momentary feeling rapidly changed from mild pleasure to absurdly furious. The person filled the information again and eventually got the tickets, yet nonetheless ended up feeling rather annoyed. Whole experience took seven minutes from start to end. Had someone collected instant utility (on a scale from –5 to 5) once per minute during her experience, reports would have been like so: 1, 2, 2, 3, 2, –4, –1. Positive values correspond to smooth part that occurred before billing process, negative values correspond to moments after site got stuck to loading screen. In overall, it seems that the experience was still quite positive ($\mu \approx 0.71$). As a sum of all the instant utilities, total utility is 5. However, according to peak-end rule, remembered utility of her experience is something quite different. When averaging most intense moment (–4) and last moment (–1) result is negative (–2.5) Peak-end rule predicts that even though during experiencing she felt mainly pleasure, yet she remembers the experience as negative. It is as past herself and present herself are in conflict with each other.

As Redelmeier and Kahneman (1996a) maintain, it is sensible to acknowledge that peak-end evaluations may lead to conflict between real-time and retrospective evaluations, though in most situations it should result in reasonable judgements. Peak-end rule holds that discrepancies occur when most intense moments and ending moments differ from the average moment of experience. If each moment pertains feelings of similar valence and intensity, there then memory should correspond with the experience.

Authors have found two consequences of peak-end effect, which are duration neglect and violations of temporary monotonicity (Kahneman et al., 1997). Not surprisingly, when people pay attention to duration, they think it as an important aspect to the experience (Kahneman et al., 1997). When a family is planning a holiday trip, they might think that they will get twice as good experience if they can extend the trip to last six days instead of three days. More time for enjoyment. Astonishingly, *duration neglect* implies that duration of an experience has little effect to remembered utility. In other words, when one contemplates how positive an experience was, how long it took does not matter. Duration neglect does not necessarily imply that durations are not remembered, it simply suggests that they are not taken into account when hedonic evaluations are formed. *Violations of temporary monotonicity* holds that remembered disutility (i.e. displeasure) of a negatively perceived experience can be reduced by adding average reducing extra period of discomfort (Kahneman et al., 1997.).

Peak-end effect and its implications have been robustly tested via several psychological experiments. Many of them has focused on negative, painful experiences. In one famous study, which consisted of two trials, participants were prompted to immerse their hand in painfully cold water (Kahneman et al., 1993). In short trial, they immersed one hand in water at 14 °C for 60 seconds. In longer trial, hand was submerged similarly at 14 °C for a minute, but additionally they held hand in water for 30 seconds longer while temperature gradually was raised to 15 °C. Afterwards, participants were asked to decide which trial

they would like to repeat. By the peak-end rule and duration neglect, majority, or 69 % (22 out of 32) of the participants opted the longer experiment. During trials researchers measured total utility by continuously prompting participants to rate the pain they experienced on a scale ranging from 0 to 14. In longer trial, most participants indicated a decline in pain of two or more points during the last 30 seconds. For them, average sum of peak-end evaluations was lower in the long trial than short trial, and consequently 17 out of 21 preferred to repeat longer trial. Results of the study suggest that in certain conditions people prefer more pain over less pain. Not always should designers try to maximize experienced utility. Another important insight gained from cold water experiment, is that when making decision participants did not spend much time in analytic comparisons — they simply opted for the memory they disliked less. Rest of the participant did not experience decrease in pain levels and majority of them preferred short trial.

In another study, which had substantial sample size (N = 682), researchers studied patients who went through rather painful colonoscopy procedure (Redelmeier & Kahneman, 1996b). Patients were assigned to two groups. In one of the groups, procedure was extended by leaving unnecessarily the colonoscope instrument in place for one minute after the actual procedure was completed. Yet less uncomfortable than actual procedure, added period was mildly painful. As with cold-pressor experiment, procedure with extended period was better perceived than shorter one. By the rule of violations of temporary monotonicity, longer procedure yielded high improvement to global evaluations of a procedure. In fact, prolonged experiment is excellent example of experience design, an adept utilization of psychological knowledge to improve user experiences (Hassenzahl, 2010). Moreover, it manifests importance of understanding sequencing effects.

Do et al. (2008) proved that judgements of pleasurable experiences are also subject to peak-end rule. In study, 104 participants were given DVD's from two lists, *A list* consisted of highly rated movies, while *B list* included only positively rated movies. Group A received DVD from only A list, Group A + B received DVD from list A and also from list B. Participants were asked to rate how pleased they were with overall experience of receiving movie gifts on a scale from 1 to 7. Results showed that mean rank for pleasure rating in group A was significantly higher than in group A + B. Objectively speaking, this implies that participants preferred less pleasure over more pleasure. A great movie is better than same movie with an addition of inferior movie. However, based on Kahneman's (2011, p. 360–362) arguments on joint evaluations, one could assume that if the same study would be carried out with within-subjects study method, A + B would be preferred over A. In similar vein, in study from the field of marketing, where study objects were television commercials, results showed that positive feelings are rated more highly if commercials had high peaks of intensity and strong positive feelings (Baumgartner, Sujan, & Padgett, 1997). Diener, Wirtz, and Oishi (2001) coined James Dean effect after findings of their study indicated that wonderful but short and abruptly ended life is perceived as better than long life of mild happiness. Hence, findings strongly support duration neglect.

Human brains are not capable of storing and retrieving every bit of data of given experience. Great deal of what we have experienced are lost and forgotten for good. As we cannot remember everything our brains have evolved to keep the most intense parts and leave the rest. Psychological studies have confirmed that this can cause biased memories. Peak-end effect states that our memories are strongly influenced by the most intense moment of the interaction and the end of the interaction. There are two interesting consequences of peak-end effect: duration neglect and violations of temporary monotonicity. Duration neglect holds that duration of an experience has little effect to retrospective evaluations of that experience. Violations of temporary monotonicity holds that displeasure of a negatively perceived experience can be reduced by adding to the end of episode an extra period of average reducing discomfort. As an outcome to peak-end effect, subsequent decisions do not correspond to experiences.

3.3.3 Prior UX research on sequencing effects

User experiences, especially ones where user interacts with a computer interface, are quite distinct from other types of experiences. Certainly, they are distinct from experiences studied in psychological experiments. Evaluating an experience of keeping hand in painfully cold water is much more straightforward than evaluating a multifaceted experience of performing multiple tasks with a visual and interactive information system. Thus, it is possible that same rules found from psychological research do not hold in HCI context. One should not absorb findings from psychology and simply apply it to UX design without applying findings in HCI research first. Nevertheless, sequencing effects have potential to play major role in formation of an experiences and are important topic for UX research (Cockburn et al., 2017).

Gutwin et al. (2016) state that when experimenting with peak-end effects, one should manipulate the sequence of momentary experiences during interaction and simultaneously keep total content or total utility of experiences equal. Secondly, retrospective experience, or remembered utility, has to be measured. Manipulating sequences is challenging in HCI studies since users can respond differently to manipulated or controlled elements (Gutwin et al., 2016). For instance, high workload tasks that agitate negative emotions to many, can bring about positive feelings of accomplishment to some participants.

Albeit being important topic, there are few studies conducted in HCI with the topic of sequencing effects and their effect on user experience. Cockburn et al. (2017) have identified, listed and summarized central findings of all five studies. Findings are presented in the table below (TABLE 3 Prior HCI research on sequencing effects). In overall, table shows that researchers have found very mixed findings among papers. This implicates that there is yet to be had a clear picture on how sequencing affects UX. More research should be carried out.

TABLE 3 Prior HCI research on sequencing effects

Study	**Topic**	**Sequence manipulation**	**Results**
Hassenzahl and Sandweg (2004)	Correlations between mental effort and perceived usability	Recency Peak Trend Peak/End	Strongest correlation between recency and perceived usability
Harrison et al. (2007)	Perception of progress bars	Various rates of progress	Significant effect of recency: accelerating progress preferred
Cockburn et al. (2015)	Preference for page sequences	Recency Peak Peak-and-End	Marginal effects of recency. None for peak; significant effect of peak-and-end
Gutwin et al. (2016)	Experience of game sequences	Peak-and-End Recency	Mixed results, depending on game
Cockburn et al. (2017)	Preference for sequences with assistive/failed-assistance	Recency Primacy	Significant effects of recency; no effects of primacy

In addition to described peak-end effect, HCI researchers have studied sequencing effects of primacy and recency effect. While *primacy effect* refers to enhanced memory and overweighed influence of initial moments of an experience, *recency effect* addresses enhanced memory and overweighed influence of last moments of an experience (Cockburn et al., 2017). It would be interesting to study what is the role of first impressions to primacy effect, as this aspect has been overlooked in prior research on sequencing effects. Also, recency effect states that the more recent the detail more easily it comes to a mind.

To present few examples, Hassenzahl and Sandweg (2004) studied how the intensity of experiences relates to summary assessments of a software's quality. 21 participants were acquired to perform series of tasks with a software tool for the configuration of hearing instruments. After each task, participant's mental effort was measured with the SMEQ questionnaire. After completing all tasks, they were asked to evaluate perceived usability of interface with 7-point semantic differentials. Internal consistency of measure is satisfactory (Cronbach's $\alpha = .85$). Results show that summary assessment of product's perceived usability is primarily based on the end of the previous experience. Intensity of mental effort measured in the end negatively correlated with perceived usability This clearly corresponds to recency effect. The more recent the detail more easily it comes to a mind. Thus, essentially researchers were able to demonstrate difference between experiencing and the retrospective summary assessment in HCI context for the first time. Interestingly enough, results also indicate that the experienced effort during the last task was the best predictor for the perceived overall usability of the product, even better than the average.

However, in unpublished follow-up study, results showed that recency effect can be counteracted by introducing a pause of ten minutes between experiencing and the summary assessment (Hassenzahl, 2010, p. 30).

Gutwin et al. (2016) studied how game player experience and its aspects are influenced by peak-end effect manipulations with two experiments. In first, scholars studied peak-end effects caused by different sequences of game difficulty. Twelve participants played two prototypical games, and twelve game sessions in each game. Among game sessions, order of difficulty sequences was manipulated while objective difficulty of sequences was identical in each sequence. In second experiment, skill-challenge balance was manipulated through peak-end sequencing. Study included initial calibration session, where participants skill level was determined. Researchers created sequences that had appropriate skill-challenge levels as positive momentary experience, and inappropriate levels as negative experience. Gutwin et al. (2016) found that in one of the three games used, game sequences with positive peaks and ends were judged as more fun, interesting, and less challenging than other sequences, despite identical total objective difficulty. All in all, despite the mixed results, authors were definitely able to demonstrate that changes in sequences can affect retrospective evaluations.

Given the shortage of research into sequencing effects in HCI, Cockburn et al. (2017) studied how recency effect and primacy effect influence user experience. They conducted two experiments with the goal of determining which sequencing effect influence the most. First experiment included 46 participants, while second 73. In experiments participants performed simple drag-and-drop tasks with assistive user interface. Researchers manipulated valences (positive-neutral-negative) of interaction sequences for a momentary stimulus. Most of the main findings originated from the second experiment. They found that positive and negative recency effect can influence user preferences, but no evidence for primacy effect was found. Furthermore, results indicate that recency effect were strongest when stimuli were recently encountered. As an important caution for future researchers, Cockburn et al. (2017) emphasise difficulty of experimentally observing sequencing effects in HCI studies, as momentary stimuli tend to be weaker than in studies outside HCI. If momentary stimulus is weak, (instant utility is close to neutral point of zero) in any given moment, unlikely there will be great variations between total utility and remembered utility of the same experience.

Five studies have conducted experiments with various products, such as configuration software for hearing instruments (Hassenzahl & Sandweg, 2004), experimental desktop software (Cockburn, Quinn & Gutwin, 2015; Cockburn et al., 2017; Harrison et al., 2007), experimental software (Cockburn, Quinn & Gutwin, 2015), and video game (Gutwin et al., 2016). Four of the objects were of prototypical nature, it means that they were not actually used in the real-world. These laboratory experiments might involve problem of ecological validity. Above all, none of the products used in studies was a website. Would be interesting to see what kind of results could be gained from experiment conducted with user-faces that actually are used in real world. Using a product that is not merely developed for the needs of experiment might cause different results.

Present thesis utilises a website that will be available for everyone with an access to web, although only in Finnish. Moreover, authors in five studies manipulated factors connected with usability side of the interaction, such as difficulty and mental effort. When measuring quality of experiencing, variables such as mental effort might be too limited for representing whole momentary experience (Hassenzahl & Sandweg, 2004). In none of the studies, momentary feelings in a pleasant-unpleasant scale were measured during the experiments.

As a summary of prior HCI studies on sequencing effects, researchers have found mixed results. These five studies investigated whether sequence manipulations during interaction could influence different aspects of UX. Results vary between studies and even between experiments pertaining to the same study. Rather than implying faultiness of theory, mixed results might be result of weak momentary stimuli and unsuccessful manipulations. Nevertheless, studies conducted so far suggest that recency effect is most prominent of the sequencing effects in UX. In each study recency effect was found in various degrees, and significant effects were found in three studies (Hassenzahl Sandweg, 2004; Harrison et al., 2007; Cockburn et al., 2015). So far, significant primacy effects have not been demonstrated, but this might be due inadequate experimental design. As of now, results indicate that designer should especially focus on ensuring that final moments of interaction go well.

Much more research needs to be done for sharpening the understanding on how events happening during experiencing lead to formation of overall UX. Many prior studies have strived to observe sequencing effects by manipulating the sequence of sub-interactions. Alike Hassenzahl and Sandweg (2004), present study aims to gain new understanding on formation of retrospective summary evaluations, not by manipulating sequences, but by examining how well momentary evaluations correspond to retrospective summary evaluations. If inconsistencies will be found, it will be interesting to see whether they can be explained with peak-end effect.

3.3.4 Experiencing versus an experience

Accordingly to Forlizzi and Battarbee (2004), distinction has been made between experiencing and an experience. Former is targeted towards present moment while latter corresponds to a memory. Evaluations given during experiencing are called real-time evaluations and evaluations given after an experience are retrospective evaluations. The question is, what is the most beneficial way to understand experiences? Should one be more concerned of experiencing or subjective memory of the experience? Which is more important, present or past? An answer depends on the perspective of the responder.

From methodological perspective, experiencing is more important (Kahneman, 1997). When trying to understand human behaviour and experiences, researcher has to gather unbiased objective data. One can select either a memory-based or a moment-based approaches for measuring experiences. By gathering data via retrospective evaluations, one might get valid data on a lucky day. Yet as studies have demonstrated, that subjects' memories of experiences are not always consistent with what actually happened or how they felt.

Memories are not exact recordings of experiential episodes, they are bits of experiences that are reconstructed every time they are fetched from the long-term memory (Norman, 2013, p. 96). Kahneman et al. (1997) explain why momentary evaluation method should be preferred over retrospective evaluations even though both methods will prevail in research:

> The advantage of real-time measures of instant utility is that they avoid the biases of memory and evaluation that affect retrospective judgments of pleasure, pain, and well-being. Because they are much easier to obtain, however, retrospective measures will remain in frequent use.

In contrast to memory-based approach, when measuring experience during experiencing with momentary evaluations, experimenter will more likely get objective data. Real-time evaluations reflect what actually happened, not what user thinks that happened.

Situation is different when one takes practitioners' or designers' point of view. Practitioners seem to prefer real-time evaluations over retrospective evaluations (Lallemand et al., 2015), but it would be wise to think otherwise. For a designer, retrospective evaluations of an experience are more important than real-time evaluations (Cockburn, 2017; Norman, 2013, p. 53) By measuring instant utility, they would get objective data, but as numerous studies have demonstrated, user might not always act in accordance to objective data. Even though momentary evaluations would yield positive total utility, product's success is not yet granted. Merely one intensively negative moment from otherwise pleasant interaction could cause user never to use a product again. That means lost revenue for the company that provides the product.

Momentary evaluations do not provide reliable way to predict success of UX in real life (Kujala, Roto, Väänänen-Vainio-Mattila, Karapanos & Sinnelä, 2011). In contrast, when remembered utility is positive, it translates to positive experience. Consequently, there is a happy customer who would likely use the product again in the future. Memories can be wrong, yet nonetheless they dictate decisions and tastes (Kahneman, 2011, p. 385). Overall impact of product is revealed only through retrospective reflection of product use (Norman, 2004, p. 88).

Experiencing to a great extent determines how an experience is remembered. Naturally, aspects pertaining to experiencing are major concern for a designer. However, significance of experiencing is indirect, whereas user's memory of experience directly influences her willingness to repeat the interaction again (Cockburn, 2017). Thus, designers should be more concerned with the memory of an experience than what actually happened. They should design for the memory of experience (Norman, 2009).

To conclude, normally experiencing trumps over an experience for UX researchers. For UX practitioners, it is other way around. In this study, we are concerned with relationship between these two types. Therefore, both are important. It is remarkably beneficial to understand how retrospective evaluations are formed in human cognitive system, since they are a matter of life and death to countless websites, online services and e-commerce sites. What happens dur-

ing experiencing does not directly matter. Instead, what is remembered from it and available to the brain's decision-making system is what counts. As Kujala et al. (2011) state, goal of UX in industry is to improve customer loyalty and satisfaction. Concentrating only on the present experience is not enough for acquiring loyal and long-term customers.

3.4 Measuring UX

Constituting of myriad of different qualities and elements, UX is rather hard phenomenon to be measured (Law, van Schaik & Roto, 2014). It has been challenging to find well-adopted method for measuring UX evaluations (Bargas-Avila & Hornbæk, 2011). Instead, numerous instruments have been created. Should UX be measured with quantitative or qualitative research methods? Bargas-Avila and Hornbæk (2011) reviewed 66 publications from 2005-2009 to see how empirical research is carried out in UX. Results indicate that UX has been measured mainly with qualitative methodologies (50% qualitative, 33% quantitative, and 17% mixed).

Most frequent data collection method is questionnaire, while many other methods have been utilised, such as interviews, user observations, diaries, and video recordings (Bargas-Avila & Hornbæk, 2011). Among 66 studies only three utilised psychophysiological measures like heart rate monitoring, galvanic skin responses. Cockburn et al. (2017) deem that psychophysiological measures have great potential to improve UX measurement in experiments regarding sequencing effects.

When assessing UX with questionnaires, in half of the articles measurements are provided with self-developed questionnaires. Unfortunately, in majority of them, questionnaire items were not provided within articles. Most frequently validated questionnaires in UX studies are AttrakDiff instrument (Hassenzahl, 2004), Lavie & Tractinsky's (2004) aesthetics assessment instrument and SAM scale for measuring emotions (Lang, 1980). AttrakDiff consists of 21 individual 7-point semantic differentials with bipolar anchors, as typical-original, cheap-valuable and technical-human. It can be used to measure products pragmatic and hedonic qualities of UX.

It might strike some surprised that among products studied in UX research, art is one of the most frequent. Other popular options are mobile phones, entertainment products such as TV, and websites. Also, imagined products are being used. Contrary to usability studies, UX studies are more interested in leisure products instead of work-related products (Bargas-Avila & Hornbæk, 2011.). There has been clear shift from desktop computing towards consumer products. Among dimensions of UX, general UX, emotions and affect, enjoyment and aesthetic appeal are the most popular (Bargas-Avila & Hornbæk, 2011).

Regarding three temporal phases of UX, scholars measure most frequently UX after interaction (71%), while researchers commonly measure experiences during interaction (58%) as well (Bargas-Avila & Hornbæk, 2011). However, UX

before interaction is widely disregarded, as only 20% of the studies have measured it. There has been lack of longitudinal studies in UX research (Bargas-Avila & Hornbæk, 2011). Fortunately, over the last decade this research gap has been noted and studies regarding how UX develops over time are increasing (e.g., Karapanos et al., 2009; Kujala et al., 2011; Harbich & Hassenzahl, 2017).

There is room for improvement among empirical UX studies. Researchers should strive for using validated methods, and if they use self-developed measures, they should at least include them in report. A study field of such a methodological variety could benefit in having stricter theoretical approach for measuring UX (Jokinen et al., 2018). However, data collection methods should be chosen according to given problem, instead of proceeding with some universal 'one for all' method (Jokinen et al., 2018). This study aims to comply with this guideline.

4 ANIMATED TRANSITIONS IN WEBSITES

In mobile applications and information visualization applications, animated transitions are well-used methods for improving UX. While such transitions are still relatively rare sight in common websites, they are increasingly apparent in modern high-end websites. When browsing award-winning websites listed in Awwwards.com, one has to exercise patience in order to encounter a website where layout transitions are not animated. This animation trend is likely to only get more popular in future. It is rather strange, however, that being such a prominent phenomenon in web, there are few studies regarding influence of animated transitions to UX (Merz, Tuch & Opwis, 2016). Hence, a website will serve as a study object in the present study. To understand potential role of animated transitions in website context, it is useful to investigate the nature, development and purpose of websites.

Firstly, historical background and evolution of websites are examined in chapter 4.1. Secondly, nature of UI animations and their role in HCI is analysed. Lastly, animated transitions are described, how they could be implemented in interfaces and what prior research says about their impact to UX.

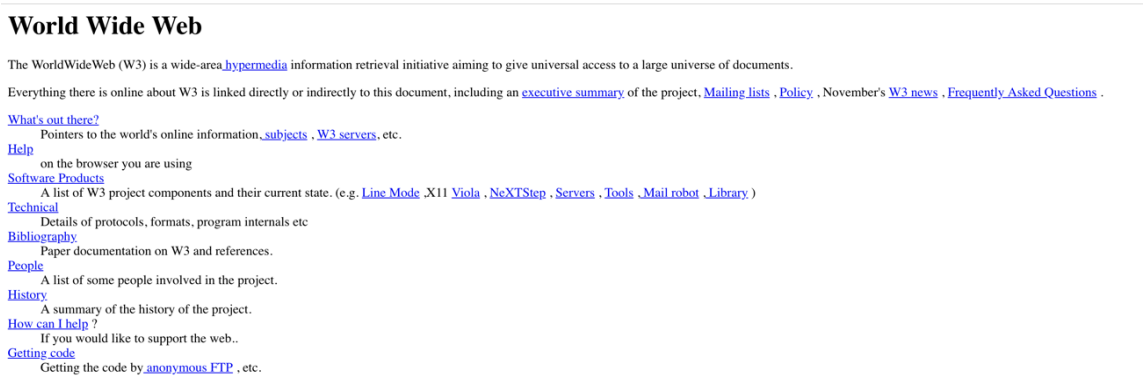
4.1 Websites

To understand websites better it is useful to take a brief look at evolution of cultural history. Collective operation of humans requires vast amount of information. For few millions of years, only place where human beings could store information was in their brains (Harari, 2014.). This restriction significantly hindered information accumulation among societies. Status quo changed around in 5000 years ago when in Mesopotamia Sumerians invented world's first information system, which enabled storing data outside brains. That is to say, writing was born. (Harari, 2014.). From that day on, there was no need to rely solely on memory anymore. Accumulation of collective information became easier, and humans became better equipped to learn from the past generations.

Fast forward to latter half of 20th century. Thousands of years after invention of writing commenced another information revolution, when humans learned to share information on an unprecedent scale with the help of digital world-wide network. The internet was born. Consequently, amount of data generated in the world exploded, and we have been able to access information regardless of place at any given time.

What are websites then? Internet consists of many different network services, and The World Wide Web (WWW) is one of the most popular means to utilise internet. The WWW is an aggregated set of every single website that can be accessed via internet. It is an information space for websites, which in turn are vessels for bringing information available for humans, and with their hyperlinks they connect every page together. Websites can be seen as a medium, that is continuation to newspapers, television, and film (Brügger, 2009). Web is a tool for collective communication. A characteristic that separates websites from other media is its interactivity (Jiang et al., 2016). Moreover, it has ability to include every other media in itself. It is certainly possible to watch television programming, listen radio and read news via website, but you cannot watch films with newspapers, for instance.

In the beginning, websites were but a hypertext documents, that could be shared by researchers spread all around the globe (Garrett, 2010). First website, published in 1991 by the inventor of web Tim Berners-Lee (Brügger, & Schroeder, 2017, p. 6), was a static document with little interactivity (figure 1). As the time went on, increasing number of users got access to internet, web technologies advanced. Websites became more and more interactive, that is, websites became able to collect information and they enabled user to manipulate information (Garrett, 2010). Evolution of web technologies has been drastic, and now they offer immense variety of tasks and actions that can be executed via the use of website interfaces from online banking to gaming. Rather than limitations of technology, human inventiveness and imagination seems to limit more what can be done with websites today. Hence, we can say that websites are much more than just one form of media.



World Wide Web

The WorldWideWeb (W3) is a wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an executive summary of the project, Mailing lists , Policy , November's W3 news , Frequently Asked Questions .

What's out there?
Pointers to the world's online information, subjects , W3 servers, etc.
Help
on the browser you are using
Software Products
A list of W3 project components and their current state. (e.g. Line Mode ,X11 Viola , NeXTStep , Servers , Tools , Mail robot , Library)
Technical
Details of protocols, formats, program internals etc
Bibliography
Paper documentation on W3 and references.
People
A list of some people involved in the project.
History
A summary of the history of the project.
How can I help ?
If you would like to support the web..
Getting code
Getting the code by anonymous FTP , etc.

FIGURE 1 The very first website, a static hypertext document (W3C, 1991)

Increased functionality seems to bring duality to the nature of website. Garrett (2010) distinguishes websites as software interfaces or websites as hypertext information page. There are highly advanced and refined functional websites, such as Google Maps and Skype, that share little semblance with first websites. They allow users to modify their contents and to change from one layout view to another. On the other hand, there are loads of basic html-sites that possess no more functionality than distribution of information. Garrett (2010) maintain that often lines are blurred, and a website contains both functional and informative elements. Before all, websites are flexible entities that both contain major part of accumulated knowledge of human kind and has successfully substituted analogous services with digital counterpart.

4.2 Animations in UI

Term 'animation' stems from Latin noun 'anima', which means life or breath. Moreover, it is closely related to verb 'animare', meaning to breath or blow; and noun 'animationem', that which is blown upon (Crafton, 2011.). The word has two meanings: it can be referred as (1) movement or (2) bringing something to life (Crafton, 2011). Nowadays, animation is presumably understood as method of manipulating inanimate images to appear as moving images. In a sense animation can be seen as creating illusion of life from lifeless objects. These illusions can be viewed and enjoyed, for instance, via television, smartphone or computer screen.

As possibilities of websites broadened, among others, one phenomenon that became prominent was animation. Websites are not only restricted to static views, text elements and pictures, there can be motion involved as well. During the evolution of World Wide Web, users have encountered all kinds of animations from GIF animations of infinitely rotating company logos (90s) to complex and heavy Flash animations of hand-drawn cartoon characters (early 00s), and more recently, modern CSS-based animations of smoothly moving graphic elements (2010s).

Fundamentally, an animation in a computer interface is a sequence of static images, with a frame rate being so rapid that human visual system cannot process what they see as separate images, and consequently an illusion of continuous visual change is created (Chevalier, Dragicevic & Franconeri, 2014). Quite similarly Bétrancourt and Tversky (2000) define animations as following (Chevalier, Riche, Plaisant, Chalbi & Hurter, 2016):

> Computer animation refers to any application which generates a series of frames, so that each frame appears as an alteration of the previous one, and where the sequence of frames is determined either by the designer or the user.

Animations in interfaces can be applied to serve copious different purposes and they can cause many different kinds of impacts to users (Chevalier et a., 2014).

Chevalier et al. (2014) identified through reviewing, building and validating taxonomies, up to 23 roles that animations can serve in user interfaces. They divided roles to five different categories, which are keeping in context, teaching aid, improving UX, data encoding and visual discourse. For instance, improving UX include roles such as keeping the user engaged; providing visual comfort and aesthetics; and revealing or hiding content. Keeping in context includes roles like staying oriented during navigation and supporting tracking during layout changes. One could ponder why UX is among the categories as does not all the categories deal with improving general UX?

If we remind ourselves with Hassenzahl & Tractinsky's tripartite definition of UX focusing around user, product and context, we see that animations are characteristic of product or system being used. Animations are one part of UI that contributes to the overall experience. They belong to both pragmatic and hedonic aspects of product (Chang & Ungar, 1995). For example, animations can be used to help user in navigating around the website, or they can be used to evoke aesthetic pleasure in user. Context and user preferences should dictate the purpose, form and style of the animation, but ultimately it is characteristic of the system.

Use of animation does not necessarily equal to improved UX. On the contrary, if not implemented with care the perceived impacts of animations can be negative (Tversky, Morrison & Betrancourt, 2002; Merz et al., 2016). Picture below represents one extreme example of bad implementation of animation in a website (Figure 1). As in example, when a web page is loaded with multiple animations that compete for the user's attention, it likely results in cognitive overload. Cognitive overload refers to situation where act of mental integration of information requires excess of cognitive resources from the working memory (Sweller, 1994). Acceptable level of distracting elements on webpages seem to depend on the culture. For instance, Asian websites tend to be complex, content-heavy and distracting, while western websites are more structured and simpler (Reinecke & Bernstein, 2011).

FIGURE 2 Example of bad use of animation (Lingscars.com)

Human eyes are biologically and evolutionarily programmed to register temporal and spatial changes in their visible environment (Johansson, 1973). Perceiving motion in environment can reveal location of prey or predator. Thus, ability of motion perception and responding to this stimulus has been absolutely vital for survival. Receiving valid information about other animal's motion is positively correlated with survival rate (Johansson, 1973). The ability does not get switched off when humans interact with technology. From otherwise static screen user cannot help but focus her attention to the element that is under state of change. This is automatic and unconscious characteristic of human cognitive system (Johansson, 1973). Now, situation is different if an interface is loaded with simultaneous animations instead of one moving element. In this situation feelings arising to user's head are likely confusion and frustration.

Essentially, animations are moving images. In order to achieve positive impact with them, animations should always serve a certain purpose. Role they serve can be pragmatic or hedonic. Additionally, they should be implemented accordingly limits of human cognition. Otherwise they just distract user from the content they are looking for.

4.3 Animated transitions

Scholars agree that one potentially positive way to utilise animations in interfaces is through animated transitions (e.g., Chang & Ungar, 1995; Thomas & Calder, 2001; Dessart, Genaro Motti & Vanderdonckt, 2011; Merz et al., 2016). Websites and other interfaces have evolved increasingly more interactive and adaptive, which means that their form and content can be modified to a greater extent. However, more often than not interface layouts are meticulously designed, but transitions between layouts less so. Sudden and surprising visual

changes between states cause confusion and difficulties to comprehend what is going on in the interface (Chang & Ungar, 1995.). Phenomenon is called cognitive destabilization when user is mentally destabilized by the abrupt confrontation of unexpected contents (Dessart et al., 2011). Causal connection between old state and new state is not clear, and user remains destabilized until she grasps the relation between past and new (Dessart et al., 2011). *Animated transitions*, which foster smooth continuous transitions between views, offer means to support screen changes (Chang & Ungar, 1995; Chevalier et al., 2014) and avoid cognitive destabilization (Dessart et al., 2011). Replacement of sudden changes with smooth animated transitions results in reduced cognitive load and ease of comprehension. Benefits can be both affective and cognitive, or pragmatic and hedonic (Chang & Ungar, 1995).

Animated transition is a particular type of interface animation, where abrupt visual change is turned into smooth one (Chevalier et al., 2014). It can be defined as agent that connects different states of UI through the use of animations (Merz et al., 2016). It is characterized by spatial and temporal parameters (Baecker & Small, 1990; Chevalier et al., 2014). In a two-dimensional interface, transition starts from one location (X_1, Y_1) and ends in another location (X_2, Y_2). Moreover, this transition from one location to another takes certain amount of time to occur. Typical duration of transition normally ranges from 300 milliseconds to few seconds, depending on the complexity of transition (Dessart, 2011). It consists of two main states, initial and final state.

Figures (FIGURE 3, FIGURE 4, FIGURE 5 and FIGURE 6) depict an example of animated transition. In the example, user opens navigation menu by clicking menu-icon of a website (Bigyouth.fr). Initial state changes to final state when blue menu element emerges from the top right corner of the page and fills the whole screen. Transition from start to end lasts less than second but duration is just enough for human eye to notice smooth continuous transition that occurs when blue small circle expands and progressively fills the page and shows its contents. It seems as contents of menu live in inside the menu icon and reveal themselves when user calls them. Assumedly, it will help users to build mental model of the website. Mental models refer to conceptual representations in people's mind that represent their understanding of how things work (Norman, 2013, p. 26).

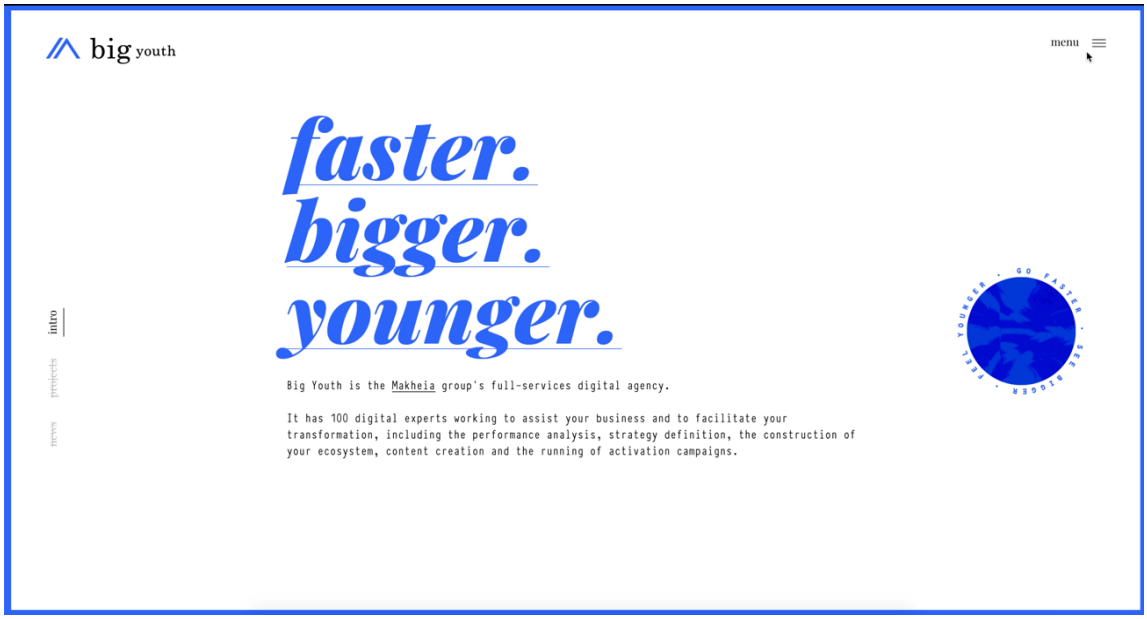


FIGURE 3 An initial state of transition (Bigyouth.fr)

FIGURE 4 Anticipation: emerged blue circle in top-right corner serves as a cue

FIGURE 5 Circle smoothly expands until it covers the whole page

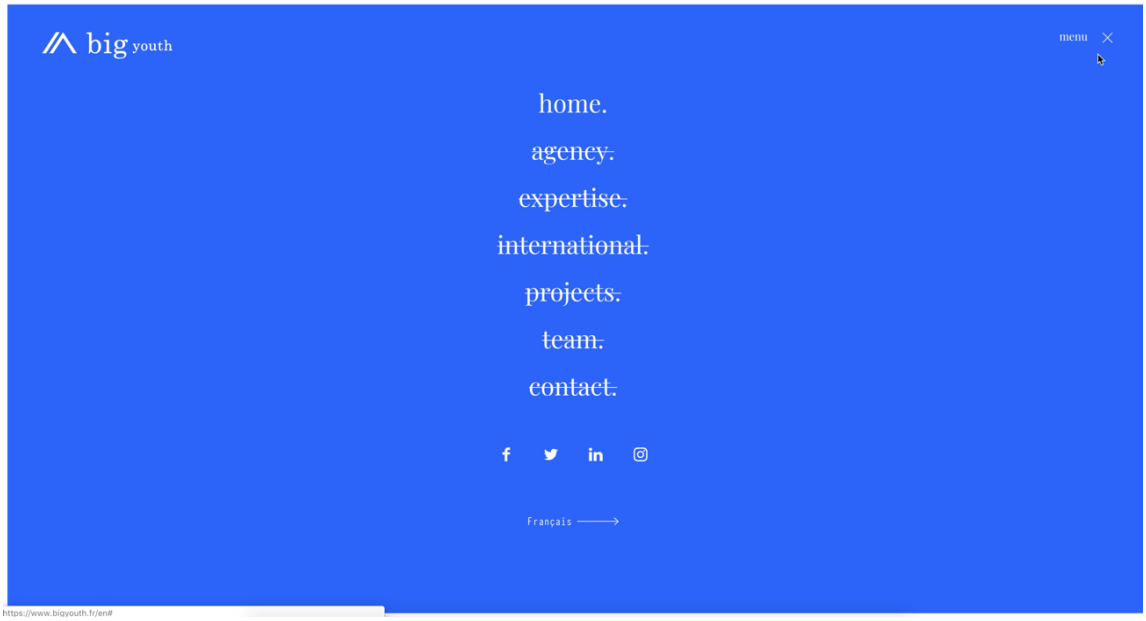


FIGURE 6 Final state of the transition

Animated transitions are inherently visual. Elements consisting of colour and form undergo change spatially in two-dimensional screen. Therefore, they possess aesthetic quality from observer's point of view. In addition to potentially making interface more comprehensible, animated transitions should have ability to arouse aesthetic pleasure to user, alike animations in general. They can be seen not only as pragmatic, but also as hedonistic element of UI. By arousing instant pleasure to user, they can provide non-instrumental value, and fulfil basic need of stimulation. Interactivity is one key factor of aesthetics (Jiang, Wang, Tan & Yu, 2016). To improve aesthetic judgements of interactions, web-

sites should be designed in a way that interaction between user and system has an interesting and changing flow (Djajadiningrat, Gaver & Fres, 2000). While interactivity and aesthetics should not be perceived as synonymous to each other (Tversky et al., 2002), one way to accomplish interesting and changing flow is through animated transitions.

4.3.1 Techniques for implementing animated transitions

Animated transitions can be implemented with various techniques. Chang and Ungar (1995) advocate applying Disney's cartoon animation principles to interface designs, as arguably appliance of cartoon techniques has potential to improve user engagement, make interfaces more pleasant and enable clear communication. Cartoon principles, which were originally identified by Thomas and Johnston (1981), can be divided to three groups namely solidity, exaggeration and reinforcement (Chang and Ungar, 1995).

Crafting an interface element to seem as solid, creates an illusion of separate, individual and interactable object. *Solidity* implies that interface elements are similar to real beings who can, for instance, move in space three-dimensionally (Chang and Ungar, 1995). It is about mimicking characteristics of real-world objects in cartoons and virtual interfaces. Solidity can be implemented through animation technique such as motion blur. *Motion blur*, or temporal aliasing, is a technique that fills the gap between old and new position resulting in illusion of continuous motion. Other technique is *arrivals and departures* - principle. According to that, animated objects should not simply materialize or disappear suddenly but maintain illusion of reality by smoothing on and off the screen (Chang and Ungar, 1995.).

To convey message more effectively, cartoons can also take liberties of what is strictly realistic (Chang and Ungar, 1995). This *exaggeration* can be used in interfaces for communicating the message to users more clearly and effectively. Like the cartoon character Road Runner spinning its legs before setting off, *anticipation* is an animation technique of exaggerating "preliminary action in order to give the audience a cue about the main action to follow" (Chang and Ungar, 1995.). It prepares user for the upcoming action (Merz et al., 2016), which can potentially decrease cognitive load of user. In the example of menu transition (Figure 4), when user hovers over the 'hamburger menu' icon, emergence of blue circle around the icon serves as a preliminary action and a cue for instructing user about the main action of transitioning menu element over current page.

Reinforcement techniques strive to convey realism in order to engage user with interface or cartoon. They can be subtle techniques that experiencer is not consciously aware of (Chang and Ungar, 1995.). *Slow in and slow out (SI/SO)* is a technique that refers to smooth acceleration of object in the beginning of motion and smooth deceleration in the end of motion (Merz et al., 2016). That is, object slowly moves out of initial state, moves faster in middle part of the movement and again slowly moves to the final state (Chang and Ungar, 1995). Another reinforcement principle is *follow-through,* addresses integration of physical law

of inertia to interface animations. It holds that when the main motion of an object ends, parts of the object keep moving little longer as a response to inertia. For example, when a person stops walking and her body stops moving, her hands will still continue moving back and forth.

4.3.2 Prior research on animated transitions' influence on UX

There has been fair amount of studies conducted regarding animated transitions among study fields such as information visualization (e.g., Chevalier et al., 2014; Chevalier et al., 2016; Heer & Robertson, 2007), human factors (e.g., Dessart et al., 2011; Huhtala et al., 2010) and HCI (e.g., Thomas & Calder, 2002; Tversky, Morrison & Betrancourt, 2002; Kraft & Hurtienne, 2017). Prior work has mainly focused on particular instances of animated transitions and found benefits in particular tasks (Chevalier et al., 2014). However, relationship of animated transitions with general UX is less studied, and it is hard to generalise those specific findings in context of websites. In web, research on animation has mainly focused on banner ads (Hong, Thong & Tam, 2007). Merz et al. (2016) state that while animations can positively affect UX, its effect depends on the purpose and animation style. Studies have found several effects of transitions, but surprisingly, there is missing evidence whether utilisation of animated transitions actually elicit positive effects that scholars have theorized (e.g., Chang and Ungar, 1995; Dessart et al., 2011).

Tversky et al. (2002) conducted literature review to test whether assumptions of potential benefits of animations hold true. Results of their study are discouraging: researchers had largely failed in demonstrating benefits of animations. Study focused on animations role in teaching complex systems. However, due the lack of prior research, review did not account for animations in the context of computer interfaces, and furthermore authors state that most promising way to use animation is especially in animating transitions.

Nevertheless, in context of computer interfaces, findings have given some support to the theory. Bederson and Boltman (1999) asked participants to navigate through interface views that utilised SI/SO styled animation transitions. They found out that animated transitions helped users to build mental maps of spatial information. They suggest that when movements are made within document, movements should be animated. Kraft and Hurtienne's (2017) studied whether transition animations in mobile interfaces positively affect formation of mental representation of the app's structure without increasing user effort. They chose between-subject study design with 35 participants per group. Their results also support finding that animated transition orientation and that they positively influence user's capability of building mental model of app's structure. UX was measured with User Experience Questionnaire (UEQ Questionnaire, 2017), and surprisingly, animations had only minor influence on UX. In information visualization context, Heer and Robertson (2007) established that animated transitions can have benefits on perception of statistical data graphics.

There is yet to have clear consensus among researchers about which animation techniques would yield the best results (Tuch et al., 2016), but SI/SO seems to stand above others. Question of which techniques to adopt of course depends on the characteristics of interface and screen size. Informal guideline for achieving the best results has been to apply SI/SO principle (Dragicevic, Bezerianos, Javed, Elmqvist & Fekete, 2011). This claim used to be given without scientific support, but there is now some evidence to support this guideline.

In late breaking study, Merz, Tuch and Opwis (2016) set out to study how animated transitions implemented with different animation principles influence perceived UX of mobile applications. They performed pilot test with 44 participants who viewed different animated transitions from their smartphones, and subsequently rated them with AttrakDiff scale (Hassenzahl, 2004). They tested transitions with three principles derived from Chang and Ungar (1995): SI/SO, anticipation and follow through. Indicative results of pilot study point that animation style of transition can affect perception of UX and animated transition using SI/SO effect reached highest means with AttrakDiff measure. In similar vein, Dragicevic et al. (2011) found that when performing accuracy tasks with desktop interface, among four different transition techniques SI/SO was by far the most accurate.

Likewise other types of animations, animated transitions need to be designed with utmost care, since inappropriately realized they can cause negative effects (Bederson and Boltman, 1999; Thomas & Calder, 2002; Dessart et al., 2011). What seems to be the most important, is that animated transitions should be used to direct focus on the most salient element of the moment. Otherwise they distract user from the main task at hand (Thomas & Calder, 2002.). Moreover, designers should ensure that duration of transition is balanced. If transition is too fast, users may not be able to make connection from old view to new view. If transition takes too long, it might cause an unpleasant lag and disrupt user flow (Bederson & Boltman, 1999). Even though duration depends on the context, Bederson and Boltman's (1999) pilot studies indicate that transitions of 0.5 – 1.0 second is optimal.

What is most interesting for the present study, use of animations in interface can cause sequencing effects. Harrison et al. (2007) demonstrated in accordance to peak-end effect that by animating progress bars in a way that they accelerate across time, progress bars were perceived as faster than they really were, and faster than other types of progress bars of same duration. Thus, accelerating animations were preferred option. Additionally, they found that pauses near the end of animation has strong negative effect on user perception. In later study, Harrison, Yeo and Hudson (2010) manipulated perceived duration with different progress bar animation designs. They demonstrated that among several options, ribbed, decelerating and backwards moving progress bar animation causes the most positive impact to user preferences, and it reduces the perceived duration by 11 %.

As with studies on sequencing effects, animated transitions have been studied with diverse set of study objects. Among others, experiments have been carried with desktop computer and experimental visual tracking software (Chevalier et al., 2014), desktop computer and Zoomable User Interface (ZUI)

(Bederson & Boltman, 1999), laptop PC with emulated mobile phone UI (Huhtala et al., 2010), smartphones with online survey (Merz et al., 2016), and smartphones with experimental mobile movie recommendation application (Kraft & Hurtienne, 2017). To author's knowledge, only one of the studies experimented with a website in their study (Merz et al., 2016), and that was only a preliminary pilot study. None have studied how transitions influence laptop website experience. This is a clear gap in research of websites. Unlike many modern websites, appearances of many experimental interfaces have been grey, dated, and in overall aesthetically rather dull.

To summarize, researchers have theorized many potential ways how utilization of animated transitions can improve UX, but certainly more research needs to be conducted to confirm these assumptions. They can improve both perceived usability and aesthetics of interface. However, present study did not find studies demonstrating animated transition's positive effect to aesthetic evaluations. Prior studies suggest that they have potential to improve experiences by engaging user and easing comprehension. Simultaneously they are hazardous tool, which, if clumsily designed, can lead to highly negative evaluations of UX (Calder & Thomas, 2002). Not many papers contain investigations of animated transitions effect to overall UX, nor in context of laptop website. Additionally, studies of Harrison et al. (2007) and Harrison et al. (2010) seem to be the only ones where topics of animation and sequencing effects converge.

5 METHOD

First part of this study was theoretical, and it was carried out with means of literary review. Review was necessary for formulating research model for the second, empirical, part of the study. In this chapter, chosen methods for the empirical part are thoroughly presented, including descriptions of empirical research approach; participants and stimuli; variables of the study; procedure of data collection and data analysis.

5.1 Empirical research approach

It came across from the prior psychological literature that people do not always evaluate experiences correspondingly their actual experiences (e.g., Kahneman et al., 1993; Redelmeier & Kahneman, 1996a). Furthermore, it was found that memories of experiences are biased towards the most intense and the last moments of the experiential episode, as predicted by peak-end rule (Kahneman et al., 1997). In addition to peak-end effects, HCI research has investigated primacy and recency effects as well. However, findings from HCI research are mixed, and there is no clear picture on how sequencing effects affect UX. State of affairs is no different with animated transitions, and we are lacking evidence regarding how beneficial they can be, how they should be implemented for maximum positive effect, or should designers ignore animated transitions altogether.

Addressing presented gaps in prior research, present thesis combines UX, biases of memory and animated transitions. First and foremost, we are interested to examine whether inconsistencies occur between real-time and retrospective evaluations of UX. Secondly, we are interested to find whether possible evaluations can be, to any degree, explained or predicted with peak-end rule, or other sequencing effects. Lastly, animated transitions' influence on UX is examined in both measurement points (during and after interaction). From the literature review we have learned that it is more challenging to demonstrate peak-end effect's influence on user preferences in HCI studies than in psychology experiments (Cockburn et al., 2017). Regardless, it is important to study se-

quencing effects. We need to gain better understanding of them as they might greatly alter user preferences (Cockburn et al., 2017).

Research questions are answered via empirical quantitative study. Quantitative method was chosen because in present study experiences are measured in a numerical form and cause-effect relationships are examined (Carr, 1994). Data is collected by conducting experimental user tests where participants interact with laptop website. Quantitative data is needed for comparing results in two different groups and measurement points. In order to measure how consistent users' memories are to corresponding experiences, one needs to collect objective data of the experience. It also asserts that data has to be collected twice: during interaction and after interaction. Kahneman et al. (1997) state that a way to measure objective data from experience is by measuring its temporal profile of instant utilities. In this context, term 'objective' refers to actual events and feelings that experiencer witnessed during experiencing. Objective data is not equal to events experiencer *thinks* that happened. Instant utility refers to what actually happened during experiencing and is not susceptible to biases of memory. In this study, instant utility is held synonymous with real-time evaluation. According to Hassenzahl's (2008) model, real-time evaluations will be collected by asking participant to rate their current affective state. Present thesis incorporates a numeric scale of instant utility ranging from –5 to 5.

As there are two measurement points, study involves two parts. In first part, user tests are conducted in controlled conditions. Participants will be asked to perform several tasks with a goal-oriented non-work-related website with a laptop. Almost simultaneously, real-time data is collected that reflects their current experience of interaction. Researcher will carefully divide tasks to different sub-interactions, so that peak-end effects can be measured by asking users to rate the current feeling at given time. Second part takes place one week after interaction. By waiting seven days, we get to measure how dynamic is UX. With survey, participants will retrospectively evaluate experience of that same interaction they had in laboratory.

As participants' affective states are measured for collecting real-time evaluations, it demands that participants are interrupted during interaction. Interruptions change the nature of experience and can negatively affect validity and reliability of collected data. Therefore, manner and timing of interruption has to be carefully planned (Bailey & Konstan, 2006). Disruptive effect can depend on the cognitive load of current task (Gillie & Broadbent, 1989) and timing of interruption in relation to phase of the current task (Czerwinski, Cutrell & Horvitz, 2000). In our study, goal was to minimise effects of interruption. Regarding timing, we chose to interrupt participants between the execution of tasks, as during task boundaries user's mental workload is decreased (Miyata & Norman, 1986; Bailey & Konstan, 2006). Regarding manner, participants were interrupted to rate their evaluations with the help of other computer, since auditory and crossmodal interruptions have been witnessed to yield more errors than other interruption strategies (Speier, Valacich & Vessey, 1999). Furthermore, tasks were not designed as very cognitively demanding to further decrease the negative effect of interruption. Unfortunately, study of Bailey & Konstan (2006) indicates that interruptions can cause negative impact to affective state and which

cannot be mitigated by decreasing mental workload. It is reasonable to expect that measurement of real-time evaluations causes, to some degree, an effect to participants' experience.

Study utilises both within-subjects and between-subjects study designs (mixed design). As a between-subjects design, participants will be divided to two groups, a control group to interact with a website without animated transition effects, experimental group with otherwise same interface with added animated effects. In order to avoid biases, apart from animated transitions two websites will be exactly identical. As a within-subject study, each participant will be questioned about their experiences twice: during the interaction, and after the interaction, as Kahneman et al. (1997) recommend. Within-subjects design is necessary for analysing differences between two measurements of one individual.

According to the peak-end rule (Kahneman, 2011, p. 380), it is expected that retrospective overall evaluations correlate with calculations of averaging most intense moment of interaction and end of the interaction. Hence, expected result to research question 1 is that there is small degree of inconsistency between users' retrospective and real-time evaluations. However, if profiles of real-time evaluations do not fluctuate, there is not much room for kind of inconsistencies that can be explained by peak-end rule. Expected result to question 2, is that users will put overweighed emphasis on these two moments of interaction. Furthermore, author assumes that expected result of question 2 explains, to a great extent, expected result of second question. What comes to third question, assumedly animated version of website will yield higher mean in UX measure derived from the AttrakDiff 2 lite measure. However, there is a chance that other hidden variables influence more than transitions, whereupon not much can be inferred from animated transitions' effect on UX.

5.2 Participants and stimuli

Initially participants (N = 25) were recruited from the mailing list for people who are interested in participating in HCI experiments. Additionally, there was an attempt to recruit participants via Facebook group of students of cognitive science. Alas, researcher did not obtain adequate resources to offer sufficient initiatives for attracting participants towards this study. Hence, altogether mailing list and Facebook group provided three participants. Thirdly, participants were recruited by asking acquaintances to participate in person. This method was proven to be the most effective, and most of the participants were found this way. Majority were mostly students or recently graduated. Five movie tickets were raffled among participants. Prerequisites for participating was that they had to be 20–40 years old, and with language proficiency of Finnish (they had to use Finnish website). Participant's average age was 26 (sd = 3.41, age range: 23–39), and 44 % of them were female. Accordingly between-subjects study design, participants were evenly distributed into two groups (animated = 12, non-animated = 13). Genders were divided equally among two groups so

that there were 5–6 females per group. Experiments were conducted in tranquil locations where participants could concentrate on their mission without interference.

As frequently mentioned above, present study is interested in people interacting with laptop websites. Therefore, it is only natural that chosen stimulus context and study object was a laptop website. The stimuli were presented to participants through Finnish website Laillisetpalvelu.fi (FIGURE 7). The website, which was built with WordPress content management system, was launched online in February 2019, only few weeks prior the experiments. Essentially, the website lists online services that offer legal content for customers. These services provide content for free time entertainment. Website's purpose is to inform, educate and change behaviour towards more ethical way of consuming digital content. It is targeted towards citizens of Finland. Listed services offer entertainment in five different categories: TV & movies, music, radio, e-books and games. The website includes a filter system that enables user to find specific services. For example, by activating two filters a user can set filters to show services that offer video game content and are free to use.

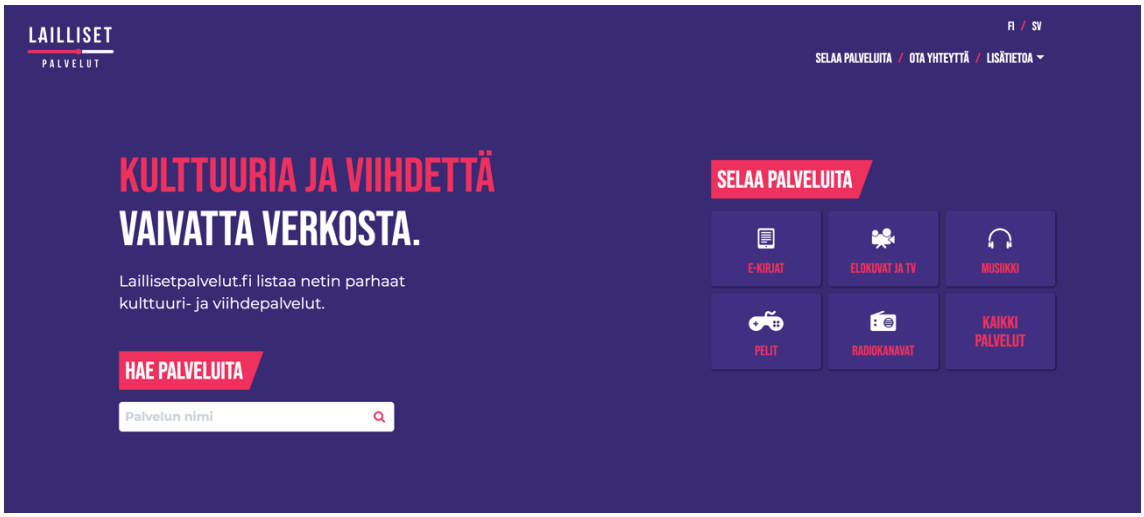


FIGURE 7 The homepage of Laillisetpalvelut.fi

Addressing psychological needs relevant to UX (TABLE 2), Laillisetpalvelut.fi can aid users in fulfilling a need of stimulation. It can be seen as instrumental means to arouse pleasure with entertainment. It does not provide pleasurable content but helps user to find it. Hence, it can be initial step in the process of consuming content. While there are many who already know where to find content they desire, not everyone knows how they can access their favourite songs or tv-series online.

Laillisetpalvelut.fi was chosen for few reasons. Firstly, the website was partly designed and solely implemented by author. This enabled author to create experimental versions of the original website by manipulating animated transitions accordingly results of literature review. Apart from transitions, experimental versions were identical with the original version. Manipulations could not have been possible in other websites where author is not a stakehold-

er in a similar position. Thus, website could be manipulated yet simultaneously used in the real world. Using a website that can be found from internet should increase the ecological validity of the study. Secondly, the website is not well known. As the website was recently launched and unfamiliar, none of the participants had used the website prior their experiment. Hence, possible biases originating from prior experiences were non-existent. Owner of the website, The Copyright Information and Anti-Piracy Centre, has granted permission to use it for purposes of research. Manipulated versions of a website used in experiments slightly varied from the original online version, as original contains both animated and non-animated transitions.

As mentioned before, manipulated website elements were exclusively animated transitions. Other than that, two versions had exactly same contents and structure. Non-animated version of website did not include a single transition that was animated. In animated version, every layout change was animated. For instance, this meant that when participant from experimental group changed a page, two big elements from the top and the bottom of the page emerged to ensure smooth transition from old page to new. This page transition is illustrated in the figure below (FIGURE 8). Clicking on menu item (1) smoothly releases two purple elements from top and bottom of the page towards vertical centre until they fill the page completely (2). Simultaneously, contents of old page fade away. Right afterwards, two elements start sliding out of the screen, to where they came from, and as they exit, contents of new page are gradually shown. Entire transition occurs in 0.8 seconds.

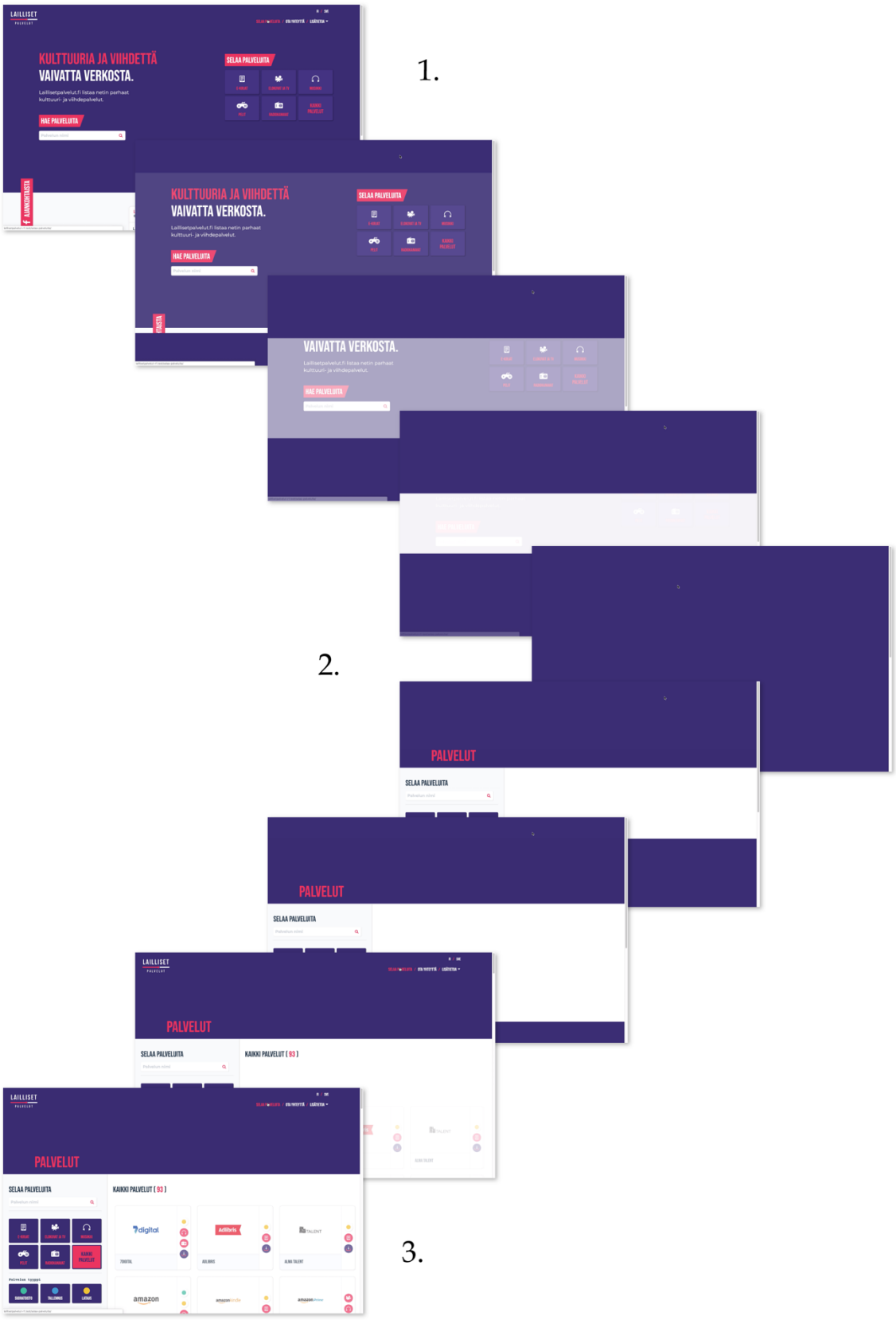


FIGURE 8 Example of animated transition utilizing SI/SO principle.

Also, when participant used filters to find different services, animated version of website did not abruptly replace visual elements representing services that did not comply with filter terms with those ones that did. Instead, unchanged elements gave room for new ones, which appeared gradually closer to the observer as if from afar. Simultaneously, elements that had to be replaced, exited the screen by gradually disappearing away from the user.

Animations adhered consistently to SI/SO principle, which was chosen as it has gained more support from the academics than other principles (e.g., Dragicevic et al., 2011; Merz et al., 2016). Animations started slowly, accelerated in middle culminating in top speed, and slowed again in the end. Animation durations depended on the size of transitioned elements and distance of transition. Durations varied from 0.2 to 1.0 seconds. Animated transitions were implemented either with Cascading Style Sheets (CSS) style sheet language, or by using JavaScript programming language. Page transition animations were done with JavaScript, but other than that, transitions were animated mostly by CSS.

Progression of experimental episode was controlled by the researcher for ensuring that each participant faced the same stimuli and in the same order. Only stimulus that varied among two groups was style of transitions. Whole interaction consisted of six sub-interactions. Each sub-interaction consisted of tasks, that had to be performed in order to continue to the next phase. If some task took inconsiderate amount of time, participant was allowed to continue without finding solution to that task. Regardless, the premise was that next task can be started only after previous had been completed. Each sub-interaction presented different parts of the website to the participants, so that combined they offered overall picture of the website. First sub-interaction consisted of getting familiar with front page and first impressions. In second sub-interaction, participants read information regarding the purpose of the website. In third sub-interaction, they got to know Facebook feed and its mechanisms. Display of Facebook posts were manipulated to show same posts for every participant. Fourth sub-interaction was about search function, and last two sub-interactions addressed the filter mechanism that enable users to examine different services listed on the site. On average, tasks were not very difficult to complete. Fifth sub-interaction included the most challenging task with the intention to evoke stronger emotions than in other sub-interactions. In turn, stronger emotions were pursued to be elicited in this sub-interaction, for it would enable calculations of peak-end effect. This was the only task where several participants gave up finding the solution. Finnish task descriptions can be found from the end of paper (appendix 1).

5.3 Variables

Stimuli was evaluated with measures that could be later separated into several dependent variables that were needed for answering the research questions. Dependent variables, which were expected to vary as a result of independent variable manipulation are listed in the table (TABLE 4) below.

TABLE 4 Variables used in research questions

Research question	Variable
Consistency between retrospective and real-time evaluations	Summary of real-time evaluations (summary RTE) immediate Retrospective Evaluation (iRE) Retrospective Evaluation (RE) Summary of sub-interaction evaluations (summary SIE)
Predictive power of sequencing effects to retrospective evaluations	iRE RE Peak-end Primacy (RTE 1) Recency (RTE 6) summary RTE
Influence of animated transition on UX	Perceived UX (part 1) Perceived UX (part 2)

Same measures were utilised in two measurement points, first one taking place during and immediately after the interaction and second measurement point taking place one week after the interaction. Although named differently, summary RTE and summary SIE employ same measures. In order to answer first question, one needed to operationalise concepts of real-time evaluation and retrospective evaluation to measurable variables. Second question involved variables of peak-end evaluation and retrospective evaluation. For additional insight, primacy and recency variables were utilised. Third question required operationalising overall UX.

Real-time evaluations (RTE) were operationalised based on Kahneman's (2000) moment-based approach, and Hassenzahl's (2008) UX model. A single momentary real-time evaluation is a measure on a 11-point Likert scale from –5 (extremely unpleasant) to 5 (extremely pleasant). Same Likert scale was used in second and third variables as well. Zero in the middle represents a neutral point. Neutral point is important as it anchors the scale, which in turn allows comparisons across persons and situations (Kahneman et al., 1997). According to Hassenzahl's (2008) model, RTE is obtained from immediate report of current subjective experience, a feeling is being influenced by a website use. While eudaimonic aspects might influence evaluations, RTE is principally a measure of hedonic experience. It measures momentary pleasure (or displeasure) derived

from the interaction. It includes information about feeling's valence (good, neutral or bad), and its intensity (from lethargic to frenetic) (Kahneman et al., 1997). With RTEs, word *feeling* is applied instead of *emotion*, because we assume that evaluations correspond to emotions that are consciously acknowledged. One cannot unconsciously pay attention to present emotions.

Summary RTE was constructed by calculating average of all instances of RTEs of single interaction. Peak-end was operationalised to measurable variable by averaging most intense real-time evaluation and the last real-time evaluation from the same RTE profile. Retrospective overall evaluation was operationalised to single number that contains participants memory of overall quality of past experience, also ranging from –5 to 5. This was collected once in the first phase right after the interaction, and twice in the second phase. Both real-time and retrospective evaluations reflect affective quality of UX.

Knowingly, unpleasant-pleasant scale equates to very simple measure of UX. Yet it should reflect the essential quality of UX. The emotion that underlies experiential episode is the most important predictor of how it will be evaluated (Hassenzahl, 2008). Positive values state that experience was perceived as good, while negative values indicates the opposite. Valence and intensity are only qualities we need to know for examining temporal progression of UX. Furthermore, because instant utility has to be measured whilst participant is interacting with a website, measure has to be as simple as possible. Answering multiple questions during experiment would distract the actual interaction too much. Reporting current feeling can be done swiftly, intuitively and without requiring much mental effort.

If one wants to gain more comprehensive picture on perceived UX, a single scale from unpleasant to pleasant is not enough. It is suitable for answering the first and second research questions, but it does not inform us which specific aspects of interaction contributed to user perceptions. As we learned in chapter two, holistic UX is a rather difficult concept to measure. In this case, it is best to go with validated and accepted measure of UX. Therefore, fourth variable, perceived UX was operationalised with Hassenzahl's (2004) AttrakDiff 2 questionnaire. Term 'perceived' was added to emphasize that measurements are derived from participant's subjective interpretations of website UX. AttrakDiff 2, being clearly more extensive than other three variables, was used for investigating influence of animated transitions to UX. AttrakDiff 2 includes two main constructs of Hassenzahl's (2008) UX model, namely pragmatic and hedonic quality. Hence, it allows more detailed comparison between two states of independent variable. To simplify procedure, lighter version of the questionnaire was selected (Hassenzahl & Monk, 2010). It consists of ten 7-point semantic differentials that represent opposites (e.g., good - bad). Contents of the questionnaire are presented at the end of the paper in Finnish (Appendix 2).

5.4 Procedure of data collection

Data collection methods should be chosen according to given problem, instead of proceeding with some universal 'one for all' method (Jokinen et al., 2018). However, when there exists validated measure for given problem, it is sensible to adopt it rather than to reinvent wheel with self-developed measure. This thesis strived to comply with these two guidelines.

Before actual data was collected, pilot studies were conducted to ensure that experiment was sufficiently designed in relation to research questions and research model. Pilot studies helped to examine whether questions answered to what they were meant to answer and whether they were adequately and understandable phrased. Additionally, pilot study helped to determine duration of single experiment, and improve and redesign tasks that participants were supposed to go through. Piloting involved two participants.

Procedure for collecting data consisted of two parts that were temporally separated from each other. In *first part*, participants interacted with a website and evaluated their experience during and after the interaction. *Second part* took place roughly one week after first, and it did not involve interaction. Instead, it consisted only of retrospective evaluation of the UX via online questionnaire. Experiment demanded two devices. Participants used website with MacBook Pro 2015 laptop with 15" 2880 x 1800 screen and Google Chrome browser. Additionally, next to the laptop they had complementary device (iPad or small laptop) for reading task instructions and for evaluating their experience. Both instructions and questions were presented with a survey by Webropol-software (webropol.fi). Complementary device enabled participants to rate their feelings without being interrupted by the researcher.

Before researcher received participants, experimental situation was set up. Browser were activated to full-screen mode, so that browser options, such as address bar, were hidden, and website itself filled every space of the screen. This was done to minimise amount of distractions for participants. Additionally, researcher filled the first two inputs of the survey himself, participant number and group number. Every second participant was positioned to control group. In order to maintain gender balance between groups, at times it was necessary to position two successive participants to the same group. Participants were not informed which version of the website they were about to use. In fact, they were not aware there existed two different versions of a website.

In the beginning of first phase, participants were briefly introduced to the study. They were vaguely informed that they took part in an UX test of a website, but they were not given more specific information regarding topics of the study. Being aware of study topics could have influenced participants' evaluations. Experiment took 15–40 minutes for each participant. Great variation in duration occurred because some participants were much more inclined to examine their feelings, preferences and website elements than others.

After introduction and before interaction, participants were prompted to evaluate their current mental state alike real-time evaluations with 11-point Likert scale ranging from –5 to 5. As real-time evaluations consist mainly of

self-assessment of current affective state, prior mental state could be assessed likewise real-time evaluations. Prior mental state was measured as it can influence forthcoming experience (Hassenzahl & Tractinsky, 2006). Afterwards, participants were instructed what they have to do during experiment. They were encouraged to remark aloud all notions that will occur to them during forthcoming experiential episode. Additionally, researcher mentioned that he will be passive during test and does not interfere with participant's experience.

When participants felt ready, they could start interacting with the website. Participants had to read out loud their current task instruction before using the website in accordance with the instructions. Each sub-interaction included 2–4 tasks. Questions involved in the tasks were answered also by speaking out loud. Researcher took notes to his notebook. Participants were allowed to progress at their own pace.

Immediately after completing each phase participants evaluated their feelings with complementary device. Participants were asked to record how they felt towards that sub-interaction they just had experienced. They did it by positioning a slider to value among the 11-point scale that corresponded with their feeling. Technically, they evaluated their past but not current experience. However, since past experience was extremely recent, it is plausible to consider evaluations as 'real time' evaluations. As there were altogether six sub-interactions, this procedure resulted in individual RTE profiles that consisted of six measurements per participant.

After having completed all sub-interactions and tasks involved, participants were asked to evaluate their overall experience with the website as retrospective overall evaluation. They evaluated it by positioning a single slider to point that corresponded with their experience. Participants did not need to use website anymore. Last task of first part was to evaluate their UX with AttrakDiff 2. Before participants took their leave, they were instructed regarding the second part of the study.

In second part, users evaluated the same website experience afterwards. Data was collected via another online survey that contained the same measures in the same order as in the first phase. When measurements are done on two different measurement points, it is important that measurements are as identical as possible. Then measurements can be regarded as comparable with each other. Nonetheless, one change was made to second measurement point: additional measurement of overall evaluation was added before sub-interaction evaluations. This was done to gather intuitive answers from the participants, that would resonate with their primary feelings towards the experience. Fear was, that if they were to evaluate sub-interactions first, more comprehensive reflection on their experience could influence their overall evaluations. Researcher took notice of dates of each participant, and when the time came, sent link to the participants via email with email addresses collected during the first part. Participants were asked to fill the survey as soon as possible. This time, participants were specifically instructed not to use the website as they filled the survey. On average, filling the survey took less than five minutes from participants.

5.5 Data analysis

Data regarding each research question were analysed with IBM SPSS Statistics program (www.ibm.com/analytics/spss-statistics-software). Data was collected with Likert scales, and semantic differentials. Hence variables were labelled and treated as ordinal in SPSS. Summary variables constructed from Likert variables were automatically transformed to and treated as scale variables by SPSS. Before analysis, data was recoded into suitable form for computer-aided calculations. Internal consistency was calculated in order to see whether variables correspond with operationalised concepts as intended. Also, correlation of variables (summary RTE, peak-end, RE, General UX) was calculated.

Distribution of means among sampling can be visualized with diagrams for perceiving overall tone of UX among participants, and for verifying whether variable distribution follows normal distribution. It is important to examine distribution as it determines which statistical tests has to be used. Examination proved that none of variables followed normal distribution, and therefore analysis was carried out with nonparametric tests (see FIGURE 9).

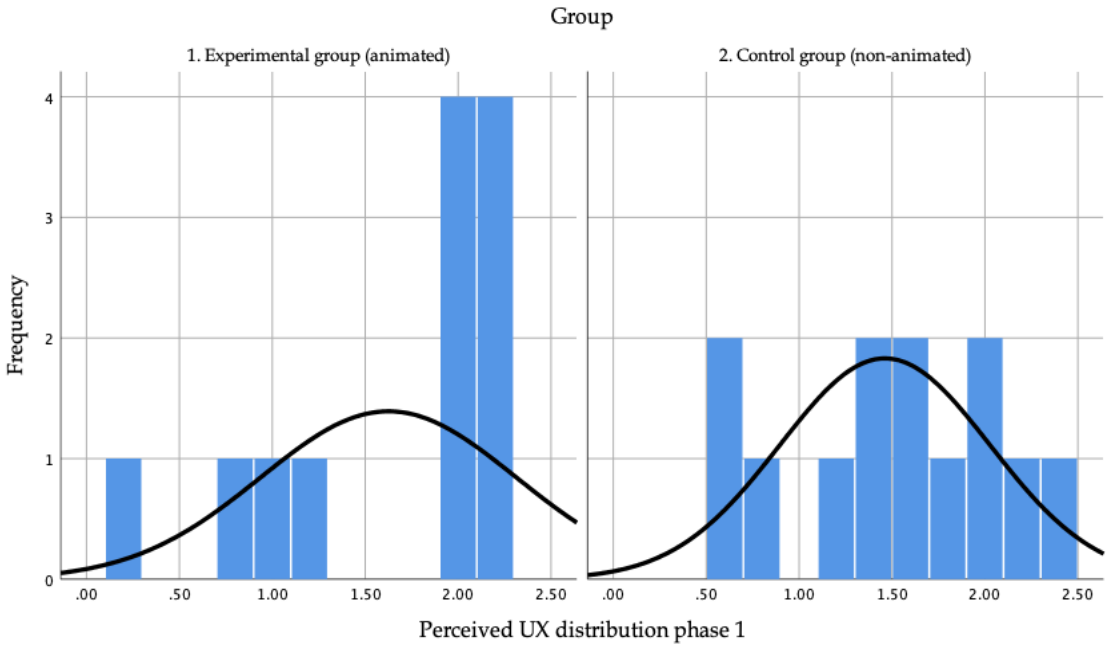


FIGURE 9 An example of means distribution among groups

Procedure of measuring momentary evaluations in first part of data collection generated temporal profiles of real-time evaluations for each participant. Similarly, procedure of measuring retrospective sub-interaction evaluations in second part of data collection generated profile of sub-interaction evaluations. Figure (FIGURE 10) shows examples of RTE and SIE profiles of two participants (P14 and P18). Numbers in Y-axis refer to values of Likert scale, while numbers

in X-axis refer to six measurements taken during episode. For P14, there are small variations between RTE and SIE evaluations while maintaining consistency in general. P18 has given more inconsistent answers than P14. In fact, for P18, there are three instances where RTE and SIE lie on opposite sides of the neutral point.

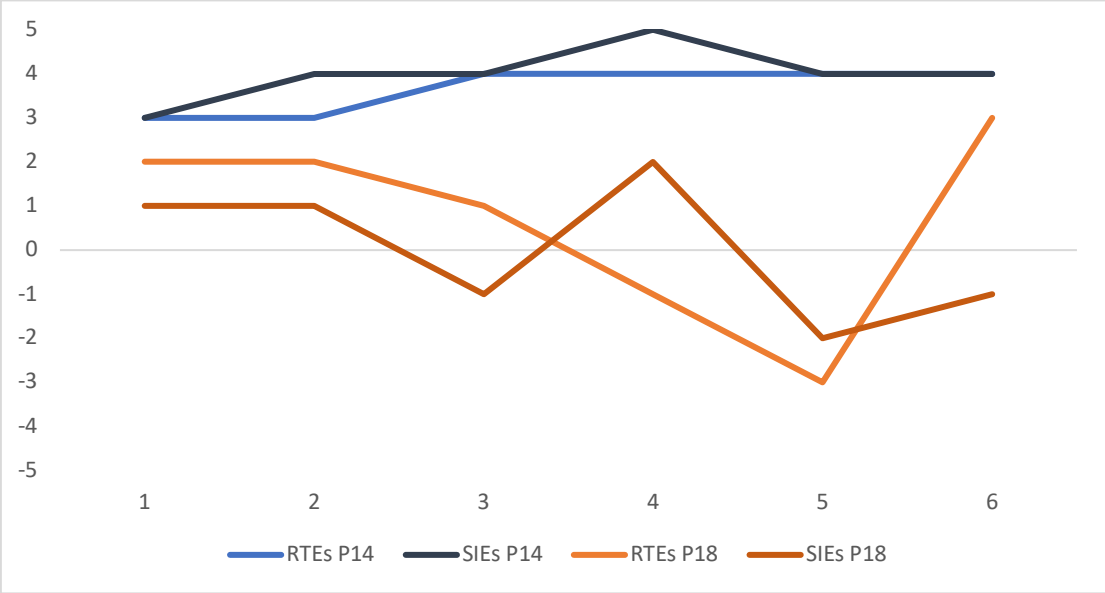


FIGURE 10 Two examples of fluctuations between RTE and SIE profiles. Likert scale can be found from Y-axis and evaluation points from X-axis.

First research questions involved comparison of real-time and retrospective evaluations. Dependent variables under scrutiny were summary RTE, iRE, RE and summary SIE. As successive measurements of dependent variables were studied, means of two related variables were compared accordingly within-subjects design. Since variable distributions did not follow normal distribution, non-parametric Wilcoxon's signed-rank test, was chosen for all three tests.

Second research question entailed investigation of sequencing effects. Primary interest was placed upon whether peak-end rule applies to retrospective evaluations. Thus, it was needed to calculate correlation between peak-end to RE. Secondarily, possible primacy and recency effects were examined by calculating two additional correlations (RTE 1 – RE, RTE 6 – RE). Additionally, correlation of summary RTE and RE was calculated. Two correlations were calculated for each variable representing momentary experience (summary RTE, peak-end, primacy, recency), because retrospective evaluations were measured twice in two measurement points. Correlations would have been calculated with Pearson's correlation if variables had followed normal distribution. As this was not the case, nonparametric Spearman's correlation was utilised instead.

For third research question, it was compared which version of website yields better UX scores with AttrakDiff 2 measure in both measurement points. Because average values were compared in two different conditions of independent variable (a website with or without transitions) and variable distribu-

tion did not follow normal distribution, nonparametric Mann-Whitney U test was be chosen.

6 RESULTS

After literature research, the following research questions remained unconfirmed: How consistent are users' hedonic retrospective evaluations to real-time evaluations? If inconsistencies occur, does the peak-end effect explain, to some degree, these inconsistencies between evaluations? Can retrospective evaluations be accurately predicted with peak-end rule? Do animated transitions influence real-time and retrospective evaluations user experience? Hence, these questions were addressed in the empirical part of the study. In this chapter, results and findings of the study are presented.

6.1 Summary variables

Table (TABLE 5) presents different variables measured during data collection. Summary RTEs were calculated for every participant by averaging all six RTE items (internal consistency was satisfactory, Cronbach's $\alpha = .87$). A week after experiment participants were asked to evaluate again each six sub-interactions (SIEs) of the experiential episode. From these evaluations, summary SIE was calculated for each participant (Cronbach's $\alpha = .92$). Peak-end variable was calculated for each participant with average of the most intensive RTE and last RTE. If variable is constructed from only two items, it is not meaningful to examine internal consistency. By calculating average of 10 items of the AttrakDiff, perceived UX variable was formed (Cronbach's $\alpha = .71$). We could assume that reliabilities of all summary variables were strong enough, and hence they were formed. Reliability of perceived UX in second measurement point was weaker than in the first measurement point.

TABLE 5 Formed variables

Variable	Number of items	Cronbach's alpha	Temporal phase
Mental state	1	-	Before interaction
Summary RTE	6	.87	During interaction
Summary SIE	6	.92	After interaction
Peak-end	2	–	During interaction
iRE	1	–	After interaction
RE	1	–	After interaction
Perceived UX (part 1)	10	.71	After interaction
Perceived UX (part 2)	10	.65	After interaction

6.2 General results

In general, and regardless of website version, UX of laillisetpalvelut.fi was perceived as positive. Pattern of positive evaluations was prominent in all summary variables and one-item retrospective evaluations. For instance, mean of summary RTE was 2.73 (N = 25, sd = 1.41, p = .000), which is quite high considering that maximum positive value is 5. Mean of iRE of first phase was even higher (μ = 3.04, sd = 1.54, p = .000), and RE in second phase almost the same (μ = 2.96, sd = 1.27, p = .000). Similarly, Attrakdiff measure yielded positive results. Means of perceived UX in two measurement points were 1.54 (sd = .62, p = .000) and 1.50 (sd = .49, p = .000). For each variable, difference of results from neutral point were statistically significant (One Sample Wilcoxon Signed Rank Test).

Before participants proceeded to interact with experimental websites, their mental state was measured in order to examine whether mental state influences interaction. Prior mental state did in fact correlate to some extent with subsequent experience. Correlation of mental state and summary RTE was statistically significant (r^s = .47, p = .019). Correlation of mental state and iRE was statistically significant as well (r^s = .40, p = .05). However, prior mental state does not correlate as strongly with RE given one week after the interaction (r^s = .35, p = .083). Checks were done to ensure that outliers did not cause stated correlations.

Results of Mann Whitney U tests indicate that males had better experience with a website than females. Mean of Summary RTE among males was 3.37 (sd = 1.24) and among females 1.92 (sd = 1.23). Mean difference of 1.45 was statisti-

cally significant (Z = –2.58, p = .010). Effect size was large (d = 1.17). Means of RE were 3.50 (sd = 1.09) among males and 2.27 (sd = 1.19) among females. Again, mean difference (1.23) was statistically significant (Z = –2.36, p = .018) and large (d = 1.07). Similarly, males (μ = 1.76, sd = .52) evaluated perceived UX in first measurement point more positively than females (μ = 1.26, sd = .65). However, mean difference of .50 was just out of the limits of statistical significance (Z = –1.93, p = .054). Yet effect size was large (d = 1.23).

6.3 Temporal profile of RTEs

Although measured with minute delay, six RTEs reflected actual momentary feelings felt during six sub-interactions of the experience. From six RTE's temporal profile could be formed for each participant. Temporal profiles illustrate affective progression of participants' subjective experience. The most negative RTE measured was –3 and most positive was 5. Hence, evaluations were quite widely distributed among scale. Altogether, 125 RTEs were measured among participants, and only 10 of them were negative while 115 were neutral or positive. As expected, RTEs measured after fifth sub-interaction contained more negative values (5) than other RTEs. Consequently, it provided the lowest mean among six sub-interactions (1.68) and highest variation (sd = 2.44). Fifth sub-interaction was intended to be more mentally challenging than others, and this proved to be true. What was unexpected, there were many participants who struggled to use the website in this part yet evaluated their experience as positive. Albeit being the most challenging moment of interaction, it was not exclusively perceived as the most intense moment of interaction. Therefore, manipulation did not work as intended. While fifth interaction was perceived as the least positive, first two sub-interactions were ranked as most positive with means of 3.28 (sd = 1.28) and 3.16. (sd = 1.59). See table (TABLE 6) below for further details.

TABLE 6 Comparison of RTEs

	RTE 1	**RTE 2**	**RT3 3**	**RTE 4**	**RTE 5**	**RTE 6**
Mean	3.28	3.16	2.52	2.72	1.68	3.04
Std. Deviation	1.28	1.60	1.50	1.88	2.45	2.01
Min value	0	–1	0	–1	–3	–2
Max value	5	5	5	5	5	5
Range	5	6	5	6	8	7
Median	3	3	3	3	3	3

After one week had passed, participants were asked to rate again each sub-interaction. These SIEs aimed to measure how accurately remembered experience corresponds with actual experience in distinct sub-interactions. Table (TABLE 7) presents results of each SIE.

TABLE 7 Comparison of SIEs

	SIE 1	**SIE 2**	**SIE 3**	**SIE 4**	**SIE 5**	**SIE 6**
Mean	2.80	2.60	2.20	2.76	2.32	2.40
Std. Deviation	1.55	1.63	1.71	1.88	2.21	1.98
Min value	–1	–1	–1	–1	–2	–2
Max value	5	5	5	5	5	5
Range	6	6	6	6	7	6
Median	3	3	2	3	3	3

With SIEs participants clearly maintain positive view of the sub-interactions, as they did with RTEs during experiencing. However, apart from fourth (SIE 4) and fifth (SIE 5) sub-interactions, remembered hedonic quality of experience had decreased in one week. In three sub-interactions (SIEs 1, 2, 6) remembered experience had decreased as much as approximately 0.5 units of measurement from the actual experience. Interestingly, most challenging sub-interaction (SIE 5), was remembered as much more pleasant than as it was experienced in the first place.

6.4 Consistency between real-time and retrospective evaluations

In this chapter, results are presented to first research question. We investigate how consistently real-time evaluations transform to retrospective evaluations. RTEs are reflected with various experiential features measured after the interaction with the help of three tests. We compare RTEs to REs taken in both measurement points. First RE (iRE) was given immediately after the interaction ended, and second REs were measured one week after the interaction. Does the six-item summary variable reflect one-item overall variables? Additionally, RTEs are compared with sub-interaction evaluations (SIEs) that were also measured after one week of interaction. Figures (FIGURE 11, FIGURE 12, FIGURE 13) show how means of summary RTE, iRE and RE are distributed over 11-point Likert scale.

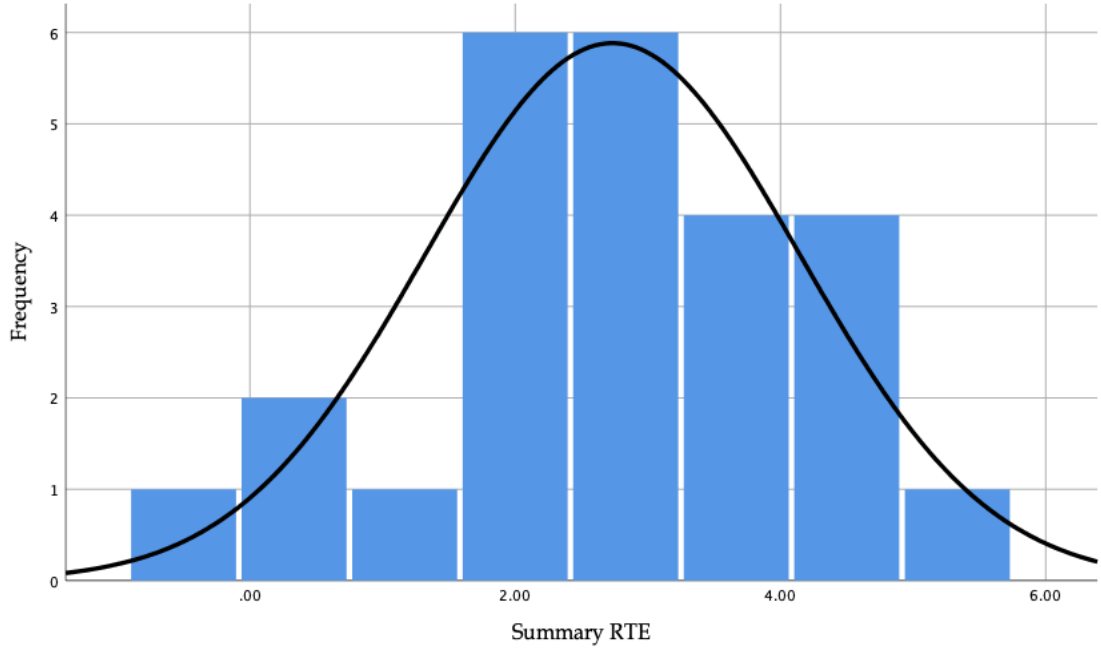


FIGURE 11 Distribution of summary RTEs

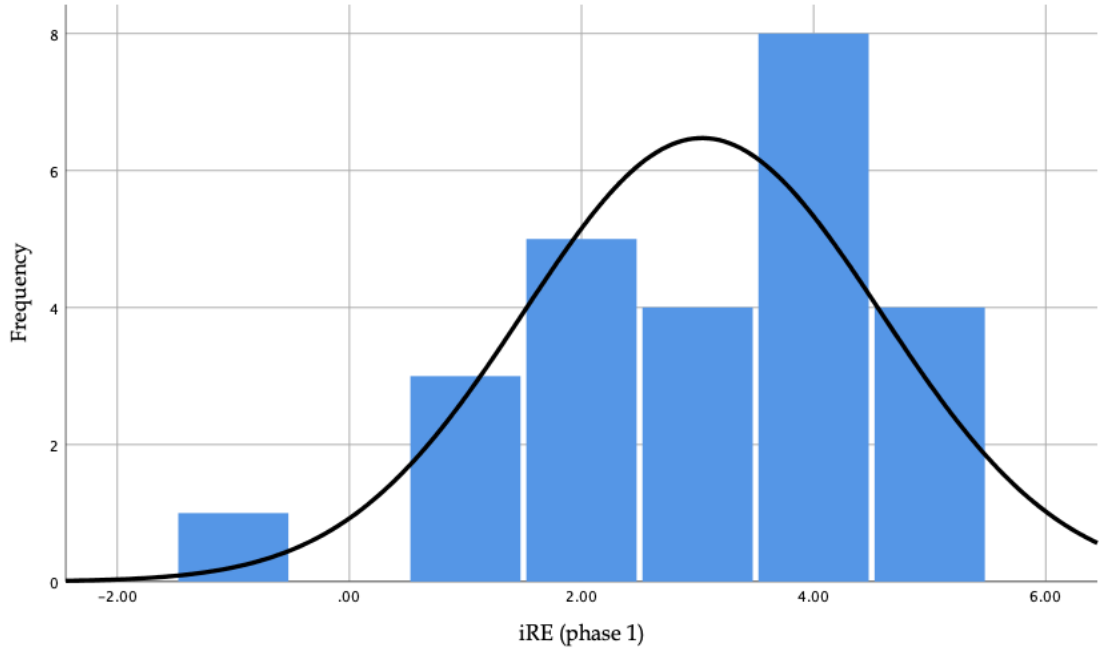


FIGURE 12 Distribution of REs given immediately after interaction

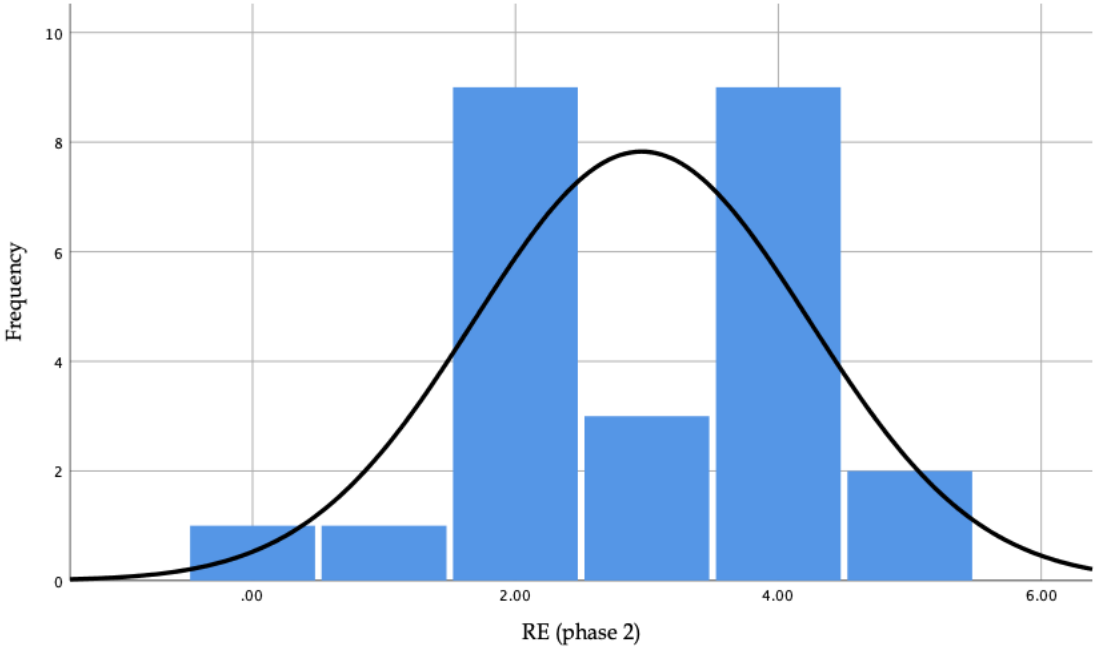


FIGURE 13 Distribution of RE given one week after interaction

As can be seen from figures above, there are slight differences in distribution of variables, although all variables are distributed almost solely over the positive end of scale. This indicates that, in general, negative experiences were not remembered as positive nor positive experiences remembered as negative. Instead, experiences were perceived as positive and they were still perceived as positive after one week had passed. From this we can conclude that there were no substantial bias of memory apparent among the evaluations of participants. As figures show, mean distributions do not follow normal distribution in any of the three variables.

Let us look at the test results related to first research question. Tests were conducted with nonparametric Wilcoxon's. Table (TABLE 8) below summarizes the results of first research question.

TABLE 8 Consistency of retrospective features to summary RTE

Variable	**Mean**	**Median**	**Std. Deviation**	**Statistical significance of difference to summary RTE**
Summary RTE	2.73	2.83	1.41	-
iRE (part 1)	3.04	3.00	1.54	No ($p > .05$)
RE (part 2)	2.96	3.00	1.27	No ($p > .05$)
Summary SIE	2.51	2.83	1.54	No ($p > .05$)

Firstly, degree of consistency between RTEs and retrospective evaluations given immediately after interaction (iRE) is analysed. Means of summary RTE and iREs were respectively 2.73 and 3.04 with mean difference of .31. Although variables differ slightly, Wilcoxon's test shows that, on average, difference between

iRE and summary RTE is not quite statistically significant (Z = –1.86, p = .063). As a result, no significant inconsistencies were found between real-time and retrospective overall evaluations given right after interaction.

Let us turn to main concern of this chapter and analyse consistency between RTEs and REs given one week after the interaction. As it turned out, results did not differ much from the previous test. Means of summary RTE and REs were respectively 2.73 (sd = 1.41) and 2.96 (sd = 1.27), with mean difference of .23. Participant's memories were slightly more positive than emotions they felt during experiencing. Mean difference was not statistically significant (Z = –1.81, p = .071). This result indicates that there were no significant inconsistencies between real-time and retrospective evaluations given week after interaction.

In second phase, participants were asked to recollect their interaction and evaluate again six sub-interactions. Mean of summary SIE was 2.51 (sd = 1.54), and mean difference between summary RTE and summary SIE was .22. While RTEs were evaluated somewhat more positive as SIEs, difference is not statistically significant (Z = –1.88, p = .061). Differences between RTEs and SIEs, as shown in tables (table 6 & TABLE 7), proved not to be that large after all. Correlation of summary RTE and summary SIE was .90 (p = .000). Thus, remembered affective experience corresponded quite accurately with actual affective experience. At least, it appears that significant memory biases were not present in participants' evaluations of distinct sub-interactions.

As a result to first research question, albeit small differences occurred, retrospective evaluations were generally fairly consistent with real-time evaluations. Even when one week had passed between interaction and evaluations, participants' recollected experience corresponded to a high degree with their actual subjective hedonic experience. An experience was thus largely in accordance with experiencing. On the other hand, it is very possible that participants remembered their experience of *evaluation* rather than their experience of interaction. Implications of results are analysed further in chapter 6.

6.5 Correlations of experiential features to retrospective evaluations

In here, answer is given to second research question. Primarily, aim was to test relationship of peak-end to retrospective evaluations. As a supplement, correlations of primacy and recency to retrospective evaluations were examined. Additionally, correlation between summary RTE to retrospective evaluations were calculated. Correlations of four variables to retrospective evaluations are presented in table (TABLE 9). Since variables did not follow normal distribution, all correlations were calculated with nonparametric Spearman's correlation. In general, variables correlated strongly with each other. Most likely high correlations can be explained by homogenic nature of evaluations. Majority of the evaluations were deemed positive, and they did not fluctuate much in the negative spectrum of the scale.

TABLE 9 Correlations (r^s) between experiential features of interaction and retrospective evaluations in two measurement points (* p < 0.05, ** p < 0.01)

Variable	**iRE**	**RE**
Summary RTE	.84**	.88**
Peak-end (*N = 23)*	.86**	.69**
Primacy	.40*	.55**
Recency	.84**	.67**

Scatterplot diagrams revealed one outlier potentially distorting results. Hence, correlations were calculated once more to ensure that outlier did not cause high correlations. Nonetheless, outlier did not change the results substantially, as can be seen from the table below (TABLE 10). Apart from primacy effect, outlier had only minor positive impact to correlations. For more information, visualizations of scatterplot diagrams can be found from appendix (Appendix 3).

TABLE 10 Correlations (r^s) without outlier (* p < 0.05, ** p < 0.01)

Variable	**iRE**	**RE**
Summary RTE (*N = 24)*	.82**	.85**
Peak-end (*N = 22)*	.84**	.64**
Primacy (*N = 24)*	.32	.49*
Recency (*N = 24)*	.82**	.63**

From the four tested variables, summary RTE correlated highest with RE given one week after interaction (r^s = .88, p = .000), and second highest with the iRE (r^s = .84, p = .000). These remarkably high correlations seem to contradict prior finding that average of momentary assessments does not reflect accurately overall evaluations of experience.

As there were two participants that experienced equally intensive positive and negative moments during interaction, peak-end number could not be calculated for these participants. Therefore, these two particular participants were not included in the Spearman's test. Test demonstrated strong correlation between Peak-end and iRE (r^s = .86, p = .000, N = 23), which was highest of all variables. Correlation of peak-end and RE was strong as well (r^s = .69, p = .000). Correlation thus weakened in the course of time. Peak-end effect to retrospective evaluations was distinctly stronger right after interaction than when measured after one week. Correlations without outlier remained strong between peak-end and iRE (r^s = .84, p = .000) and between peak-end and RE (r^s = .64, p = .001).

First evaluation of experiential episode (RTE 1) represents primacy variable. In both measurement points, correlation of RTE 1 was the weakest. Even so, correlation of RTE 1 was relatively significant with iRE (r^s = .40, p = .045) and even higher with RE (r^s = .55, p = .004). Correlation thus strengthened over time. However, correlation between RTE 1 and iRE without outlier (N = 24) were distinctly weakened (r^s = .32, p = .126), and less so between RTE 1 and RE (r^s = .49, p = .015).

Last evaluation of experiential episode (RTE 6) represents recency variable. Correlation of RTE 6 with iRE was very strong (r^s = .84, p = .000). After one week had passed, correlation was slightly reduced (r^s = .67, p = .000). Alike peak-end effect, recency effect was stronger when measured right after interaction than when measured after one week. Yet correlation was very high in both measurement points. Correlations without outlier remained strong between RTE 6 and iRE (r^s = .82, p = .000) and between RTE 6 and RE (r^s = .63, p = .001).

When correlation is over .50, one can say that effect size is large. Thus, effect sizes of summary RTE, recency and peak-end were very large in both measurement points of retrospective evaluations. As correlations of those three variables were almost equal, one cannot say any distinct experiential feature that primarily explained retrospective evaluations in this study. Primacy effect's effect size was large in second measurement, but medium in the first measurement point. It is possible that increased correlation from first measurement to second can be explained by pure chance.

As a result to second research question, peak-end predicted quite correctly retrospective evaluations. But primacy and summary RTE predicted to equal or better degree. In contrast to prior research, real-time evaluations correlated even higher than peak-end with retrospective evaluations. Nevertheless, since subchapter 6.3 demonstrated that study did not find significant inconsistencies between real-time evaluations and retrospective evaluations, these findings are less interesting. Furthermore, high correlations are no surprise when evaluations tended to be homogenously positive and variance was rather small. Regardless, it is interesting that merely a single temporary evaluation (RTE 6) can predict retrospective overall evaluations of UX nearly as correctly as average of all temporary evaluations.

6.6 Influence of animated transitions on evaluations

In here, animated transitions are under examination, and results to third research question are presented. As sample size was rather limited (12–13 participants per group), it is important to note that one should not make far-fetched conclusions from results of this research question. Results, which are summarised in table (TABLE 11), showed that there were no great differences between experimental and control group. In general, use of animated transitions did not improve perceived UX among participants. For what it is worth, animated transitions did not decrease perceived UX neither. Differences between groups were also examined with variables of summary RTE (Z = –.87, p = .383), iRE (Z = –.31, p = .759) and RE (Z = –.97, p = .331), but in none of the cases difference was statistically significant.

TABLE 11 Difference of perceived UX between experimental and control group

Variable	**Mean of experimental group (animated version)**	**Mean of control group (non-animated version)**	**Mean difference**	**Mann Whitney U test (Z)**
Perceived UX (part 1)	1.63 (sd = .69)	1.46 (sd = .57)	.17	–.98
Perceived UX (part 2)	1.48 (sd = .47)	1.52 (sd = .52)	.06	–.14

Perceived UX was measured twice, firstly a minute or so after interaction (first measurement point) and secondly one week after interaction (second measurement point). Alike RTEs and REs, perceived UX was generally perceived as positive among both groups. Means of perceived UX in experimental and control group were respectively 1.63 (sd = .69) and 1.46 (sd = .57) when measured right after interaction. Mean difference was .17. Mann Whitney U test demonstrates that difference between groups was not statistically significant (Z = –.98, p = .325).

After one week had passed, perceived UX had slightly declined in experimental groups while it had slightly increased in control group. Means of experimental and control group were respectively 1.48 (sd = .47) and 1.52 (sd = .52) when measured one week after interaction. When difference between groups is .04, it is unsurprising that statistical significance was not found (Z = –.14, p = .892).

As a curiosity, males of experimental group (N = 7) rated higher means for summary RTE, iRE and RE than males of control group (N = 7). Means of experimental group were 3.98 (summary RTE), 4.00 (iRE) and 4.14 (RE). In contrast, means of control group were 2.76 (summary RTE), 3.00 (iRE) and 2.86 (RE). Mean difference was statistically significant with RE (Z = –2.15, p = .032), and nearly significant with summary RTE (Z = –1.79, p = .073) and iRE (Z = –1.92, p = .056). Perceived UX in both measurement points was higher with experimental group (μ = 1.91, μ = 1.73) than control group (μ = 1.60, μ = 1.65), but differences were not even close to significant (Z = –0.90, p = .366 and Z = –0.32, p = .747).

As a results to third research question, it is clear that animated transitions did not greatly influence perceived UX among participants. The effect was insignificant in both real-time and retrospective evaluations. Among male participants, website with animated transitions was perceived as more pleasant than version without transition. However, with sample size so minuscule, one cannot make conclusions much from this result. It seems animated transitions are rather challenging study topic. Challenges and difficulties pertaining to animated transitions are more elaborately discussed in chapters 6 and 7.

7 DISCUSSIONS

Now that results are presented above, it is appropriate to consider more closely what they mean and how they fit into the picture shaped by prior research. Present chapter begins by addressing research questions and continues to consider implications of findings to researchers and practitioners. Lastly, methodical contribution is discussed.

7.1 Answers to research questions

Primary objective of this study was to investigate how accurately an experience corresponds with experiencing. Or more precisely put, how well users' evaluations given after the experience correspond to feelings actually felt during the experiencing when experiential episode consisted of interacting with a laptop website. The objective was investigated with the following research question:

- *How consistent are hedonic retrospective UX website evaluations to real-time evaluations?*

As it turned out, retrospective evaluations were generally fairly consistent with real-time evaluations. Retrospective overall evaluations were measured twice, immediately after the interaction ended and again after one week had passed. Participants perceived overall experience as slightly more pleasant than they perceived distinct sub-interactions on average. Difference was small, and evaluations were consistent in both measurement points. Other studies on sequencing effects have strived to manipulate experimental situation in a way that some sequences of experiential episode would be perceived as more negative or positive than other sequences. With the exception of fifth sub-interaction, this study did not involve such sequence manipulations. In general, stimuli manipulation has been difficult in HCI studies, and neither this study was successful in generating desired effect with manipulations. Aim was to arouse negative feelings by increasing level of challenge in sub-interaction. Unfortunately, in this case

increased level of challenge did not correlate with negative feelings. As the sample size was small, findings of first research question cannot be generalised to population without hesitation.

In order to further analyse findings of first research question, second research question was formulated as following:

- *Can retrospective overall evaluations be predicted with peak-end rule (and other sequencing effects)?*

Kahneman et al. (1997) formulated peak-end rule as result to psychological studies regarding experienced utility. Peak-end rule was formulated for the reason of average of the most intensive and ending moments of experiential episode correlated highly with retrospective evaluations of episode. In other words, peak-end rule predicts the hedonic quality of the remembered episode. In present study, peak-end correlated highly as well with retrospective evaluations. Thus, among this particular sample, peak-end rule could be used for predicting relatively accurately how experience was remembered. Based on results, one cannot claim that theory by Kahneman et al (1997) is always correct. However, results were not certainly in opposition with the theory. While, peak-end correlations were high, correlations were very high with other variables measured during experiencing as well. Predictive power of summary real-time evaluations was even higher than that of peak-end. Last evaluation correlated with retrospective evaluations nearly identically to peak-end. In both cases, correlation decreased over the course of seven days, yet maintaining strong levels. Retrospective evaluations could not be predicted with primacy variable as accurately as with peak-end and recency variables.

Third and last research question was a bit divergent with two previous questions and focused on animated transitions. Third question was:

- *How animated transitions influence UX right after interaction and one week after interaction?*

Present study did not find significant effects of animated transitions to perceived UX. This indicates that animated transitions did not influence UX in any way, positively or negatively. Means of perceived UX were surprisingly equal among experimental and control groups, and they were consistent in both measurement points. Moreover, animated transitions did not influence consistency of retrospective evaluations. For obtaining more robust results, one should repeat the study with much larger sample size. Nevertheless, within-subjects design, where each participant compares both versions of website, might yield more valid results than between-subjects study design.

Participants were not told beforehand that effect of animated transitions is being studied in the experiment. In that regard, aim was to mimic conditions of the real world, where users as autonomous persons are not told beforehand where to turn their attention to. Results indicate that, at least subconsciously animated transitions do not influence perceived UX. Majority of participants of experimental group did not seem to consciously notice existence of animations.

This could be explained with phenomenon of selective attention: since humans are bombarded with incessant flow of sensory information, they are not capable of paying attention to every bit of information. Rather, they focus on elements deemed as most important while the rest are left out to background, beyond the reach of consciousness (Lavie, Hirst, De Fockert & Viding, 2004.). Perhaps, as participants were more concerned with the tasks they were asked to perform, they did not have mental capacity to pay attention to transitions. None of the participants who used non-animated website did not specifically claim that website would be better with animated transitions.

Results might have been different, had researcher specifically asked participants to pay attention to transitions. Most of the participants belonging to experimental group did not mention animated transitions during interaction. Those who did (in fact, researcher recorded to notebook three occurrences where participants consciously addressed the topic), did not become irritated by the stimuli, rather they were pleasantly surprised. For one participant, appearance of animated page transition evoked emotions of surprise. Alas, researcher could not tell whether surprise was pleasant or not. Nevertheless, it is possible, that if everyone had paid attention to transitive elements, they would have perceived website as more pleasant.

7.2 Theoretical contribution

Undoubtedly this is very first HCI study that has simultaneously tackled in a such way both temporal aspects of UX and animated transitions. Perhaps, this should remain as an only study that did so. Both topics are large enough to deserve their own branches of literature, but nonetheless this study presents interesting yet slightly over-ambitious endeavour to combine those two topics for the means of master thesis. It is quite rare that animated transitions are studied in HCI, and especially when their effect to perceptions is studied over time.

It has been proved that sometimes memories do not reflect very well actual experiences (e.g., Kahneman et al., 1993; Redelmeier & Kahneman, 1996b; Do et al., 2008). Moreover, similar memory bias has been found to occur in specific HCI context as well (e.g., Hassenzahl & Sandweg, 2004; Harrison et al., 2007). Kujala et al. (2011) state that "evaluating momentary user experience is in most cases not very reliable for predicting user experience". It is possible that episodes experienced as very unpleasant or even painful can be later remembered as significantly less unpleasant than they really were. Similarly, experiences felt as positive during experiencing can be later found much more negative than they were. In contrast to these possibilities, memories were consistent with the actual experience in present study.

However, results do not necessarily contradict theory of experienced utility (Kahneman et al., 1997). It is not hard to find examples, where memory can be consistent with actuality, even though memory would be largely based on the peak and ending moment of an episode. Perhaps biases occur only under certain circumstances, and experimental design used in this study did not pro-

vide these circumstances to participants. At least, an experiential episode should contain some degree of variance in intensity and valence of feelings.

Two possible explanations for the results emerges: either (1.) participants did not forget their evaluations, or (2.) they did not forget their experiences. Did retrospectively given evaluations measure experience of evaluation or experience of interaction? Former explanation would be more desirable, as otherwise collected data would not correspond to research problem. This ecological fallacy would starkly decrease validity of the results. Questions were phrased in a manner to reduce possibility of second explanation. They were specifically targeted towards the experiential interaction, not the evaluations. Nonetheless, there is a chance that participants recollected their evaluations instead of experience.

Kahneman (2011, p. 96) demonstrates that when faced with cognitively challenging question, people sometimes provide an answer with principle of substitution: they unconsciously find related but easier question, namely a heuristic question, and answer that in place of the original question. In this study, it could have been that substitution occurred with the retrospective evaluations. Retrospective evaluations are mentally more challenging to perform than real-time evaluations. It can be difficult to make experiential synthesis of certain sub-interaction, let alone of the whole interaction, if the recollected memory is blurry, for instance. In momentary evaluations, one only needs to translate her present feelings to corresponding number in the scale. Addressing level of difficulty, there is also difference between producing overall evaluation right after interaction and one week after evaluation. Considering this, participants were instructed to answer intuitively without too much thinking. Yet, it could have occurred with retrospective evaluations that satisfactory answer did not emerge, which made them answer with the heuristic question of recalling the numbers they gave during real-time evaluations. Recalling numbers should be easier than recalling and analysing experience.

It is also possible, that participants answered according to second explanation. Real-time evaluations are easier to produce, but retrospective evaluations are only slightly more difficult. In that case, there is no need to answer by substituting the real question with heuristic question. They could have made the effort of recollecting their memory of interactive episode. If assumed that second explanation explains the results, participants' memories were largely unbiased. In 2.3.4, importance of actuality was compared to memory, and conclusion was made that memory is more important than actuality (Norman, 2013). What user thinks after interaction is more relevant than what user thought during interaction. Results of this study imply that this comparison is not very meaningful, since experiencing corresponds accurately to an experience. When one measures one type of experience, she will know the other. However, notion should be cautiously applied only to positive experiences where momentary feelings do not fluctuate much towards negative at any point of temporal profile.

It is not out of question, that there is a third explanation for how participants formed retrospective evaluations. Perhaps participants did not even make the effort of retrieving exact facts from their memories. Instead, they evaluated

based on what could have been be true. Reder (1982) states that people often recall memories with strategy of plausibility judgement: they recall by judging what could have plausibly happened, not by actually retrieving information from the memory.

Results of the study support Hassenzahl's (2008) view that momentary feelings of pleasure or pain truly are in the centre of experience. Momentary feelings measured during experiencing correlated highly with retrospective evaluations, and correlations were high even after one week had passed. As chapter 2 demonstrated, there are many other aspects to UX than mere feelings. However, it appears that affective measurement of UX gives us good overview of the overall quality of interaction. In this manner, study contributes to ongoing attempt to find key aspects of UX.

One peculiar observation was made during experiments. Namely, a service might not necessarily need to fulfil one's *own* needs in order to be positively perceived. A user can perceive website as positive because she feels that it can be useful for someone else. In experiment, a participant admitted that this website is not for her, but she can see why others could use it. Therefore, it is a good website. She might not want to visit website again yet evaluated her experience with it as positive. Without more valid evidence, not too much should be concluded from these remarks, other than that they are interesting thoughts that could be further studied in co-experience research.

Table (TABLE 12) presents summary of HCI studies addressing sequencing effects updated with findings of present study.

TABLE 12 Updated summary of HCI research on sequencing effects

Study	**Topic**	**Sequence manipulation**	**Results**
Hassenzahl and Sandweg (2004)	Correlations between mental effort and perceived usability	Recency Peak Trend Peak/End	Strongest correlation between recency and perceived usability
Harrison et al. (2007)	Perception of progress bars	Various rates of progress	Significant effect of recency: accelerating progress preferred
Cockburn et al. (2015)	Preference for page sequences	Recency Peak Peak-and-End	Marginal effects of recency. None for peak; significant effect of peak-and-end
Gutwin et al. (2016)	Experience of game sequences	Peak-and-End Recency	Mixed results, depending on game
Cockburn et al. (2017)	Preference for sequences with assistive/failed-assistance	Recency Primacy	Significant effects of recency; no effects of primacy
This paper	Consistency of real-time and retrospective evaluations	Peak	Strong correlations for recency and peak-end

In prior HCI studies, recency effect has tended to be stronger than peak-end effect. Lack of primacy effect among prior literature is quite remarkable considering how important role first impressions should have in UX (Lindgaard et al., 2006). Findings of this study are consistent with prior HCI literature in a way that peak-end and recency effects were found stronger than primacy effect. However, after one week, correlations between first real-time evaluation and retrospective overall evaluation were not considerably lower than recency and peak-end. It would be interesting to know did primacy correlation increase over time because of random variation or did it happen as importance of beginning in experiential episode tends to increase over time. Considering small sample size, former option seems more likely.

Among participants, there were three experiences where strong sequencing effects could have occurred. This was due to the fact that three participants felt rather intense negative momentary feelings (RTE 5 = –3) towards some sub-interactions, while for the most parts the experience was deemed as positive. However, those rather intense negative feelings did not distort subsequent overall evaluations to negative. This does not contract peak-end rule, as each of three participants felt equally or more intense positive momentary feelings during interaction.

Concept of satisficing might explain to some degree surprisingly small differences of perceived UX between control and experimental group. Satisficing entails that due to limited cognitive capabilities, people do not seek optimal performance. Instead, they satisfice to first acceptable alternative that comes to their mind (Oppenheimer, Meyvis & Davidenko, 2009.). Perhaps participants did not expect optimal UX from the site. Although lacking in smoothness, non-animated website being good-enough was sufficient to generate positive evaluations.

There exists scant research on animated transitions' effect to UX. Prior research has found but a minor effect (Kraft & Hurtienne, 2017). Merz et al. (2016), found that SI/SO principle was perceived as better than other two animation principles, but they did not study how much better interface was with SI/SO than without any animation. Moreover, experiments were conducted with mobile phones. To our best knowledge, there is yet to have properly conducted study in context of laptop websites. Alas, even by the present thesis, state of affairs has not changed for much better.

7.3 Design contribution

Prior papers have advised practitioners to prioritise memory over actuality (Cockburn, 2017; Norman, 2013, p. 53; Norman, 2009). It means that UX designers should design products in a way that maximises remembered quality of UX. It does not matter if one does not feel pleasure during experiencing as long as that same experience will be remembered as positive afterwards. Redelmeier and Kahneman (1996b) successfully utilised their insights on memory biases,

and were able to improve patients' experiential episode by adding unpleasant episode in the end.

Designing experience of medical procedure can be quite different from designing UX for websites. How does one design for the memory of interactive episode? At which point of the episode should one place special interest to gain best results? There are no simple answers to these questions. While consensus is yet far away, prior research suggests that most importantly practitioners should focus on the end of interaction (Hassenzahl & Sandweg, 2007; Cockburn et al., 2017). It appears that first moments are less important.

Designers cannot always control how interactions progress with websites. Commonly, it is assumed that front page is starting point of interaction with websites. Yet it is not rare that visitors arrive to certain website from some other website via a hyperlink, that directs visitor to some other page than home page. Fortunately, there is much that can be achieved by good design — even control of user progression. For example, designers can collect data regarding where users commonly begin their interaction and what is the page or element that is most likely last viewed last. Which is more common, users landing to home page or blog page? When it is known which are most likely going to be the first, the last or the most intensive sub-interactions with a website, designer can start paying extra attention to those parts of the website.

Results of this study, which should be taken with a grain of salt, indicate that a designer need not worry retrospective aspects of experience since they perceptions occurred during experiencing corresponds so well to retrospective perceptions. But this finding should be applied only to kind of positive experiences where momentary feelings do not fluctuate much towards negative. This study suggests as well that more emphasis should be placed upon ending moment than first moment of interaction.

This study cannot say much regarding which animation principles should be preferred over others, since only one principle was utilised in experiment (SI/SO). Yet there is something to be said relating to use of animated transitions in general. Based on findings, it is not uncertain that good UX can be achieved without any use of animations nor animated transitions. It can be claimed that animated transitions are not priority number one in interfaces. Website version without animated transitions generated very positive UX evaluations (approx. 1.5 out of 3), and they remained positive even after time period of one week. On the other hand, transitions, which were implemented with SI/SO principle, did not decrease perceived UX. In many participants, presence of animations was not simply consciously noticed. Thus, a practitioner could argue that what is the point of implementing them in the first place.

When resources, especially time, are limited, it might be best to opt for implementing website without animated transitions. Implementation process can be time consuming and difficult, and there is possibility that results are only detrimental to perceived UX, especially if there is not enough time to design them properly (Bederson and Boltman, 1999; Thomas & Calder, 2002; Dessart et al., 2011). But in the future, implementation of animations can get easier, and already there exists animation tools that can be easily applied to web development. Importantly, much research needs to be conducted in order ascertain pos-

sible benefits of animated transitions to UX. It is yet very much possible that animated transitions can have positive impact to UX. Otherwise practitioners might have already ceased implementing them to interfaces. They can elevate good website to another level. If future studies were able to demonstrate positive effects of animated transitions, they would provide flexible opportunity for practitioners to achieve superb UX and competitive edge.

7.4 Methodical contribution

Issues discussed in 7.2. were also closely related to topic of methodical implications, but here matters relating to methodical issues are specifically targeted. A simple scale from unpleasant to pleasant, derived from Hassenzahl (2008) and Kahneman et al. (1997) was found to provide successful way of measuring momentary hedonic feelings of experiential episode. This measure did accomplish the challenge of capturing essential information from the experiences. As mentioned, momentary feelings measured during experiencing highly correlated with retrospective evaluations and with validated AttrakDiff measure as well. When evaluating with such scale and slider, measurement takes little time and its unwanted interruptive effect from the interaction is small. It demonstrates that even remarkably simple measures can be useful in such a complicated field of UX. Unpleasant-pleasant scale can be used in occasions where it is not possible measure UX with lengthy and comprehensive measurement tools or when it is sufficient to merely measure overall valence and intensity of the interaction.

In this study, data was collected in each temporal phase of the interaction (before, during, after), which is achievement where majority of UX studies fail (Bargas-Avila & Hornbæk, 2011). Acknowledging all temporal phases is important as events occurring in each phase can alter perceptions. It is noteworthy that only one item, which represented prior mental state, was collected before interaction. More interest was placed upon other two temporal phases.

Cockburn et al. (2017) state that in order to study sequencing effects, one needs to manipulate sequences of interaction. To great extent, this study did not apply with this notion, because intention was to primarily study how perceptions of interactions evolve. In this way, present study diverges from other sequencing studies. Interest lied in investigating whether sequencing effects, especially peak-end effect, occur in common everyday interactions.

Addressing third research question, it was found that selected research model was not adequate enough for properly investigating effects of animated transitions to UX. At least, it was not suitable for studies with such limited sample size as in present study. If one wanted to study influence of animated transitions on UX with between-subjects design and in situations where users are not aware of the research topic, there should be manifold participants per group. If large sample size cannot be acquired, it would be sensible to choose within-subject over between-subjects design.

It is difficult to design between-subjects experiment in a way that influence of transitions can be observed. There is a possibility that when animated

transitions are well designed, participants do not notice them. Moreover, when comparing differences of two means with between-subjects design, participants do not have an anchor where they can base their evaluations. This can lead to issues, such as satisficing as explored in 6.2. If participants were specifically instructed to pay attention to transitions, problems such as one arising from selective attention would be eliminated. When each participant sees both versions of the website, it is easier to observe and measure whether they find animated transitions as pleasant or not. But if users, in general, do not even notice existence of transitions, as it appeared in present study, does it matter whether they like animated transitions or not? Arguably yes, because users do not need to know what *it* is exactly that causes them to perceive products as positive, as long as *it* does the trick. One possible methodical option for studying transitional effects would be to show both versions of website to participants without explaining beforehand how versions differ from each other.

Then again, even if study is conducted with between-subjects design and involves an anchor, as non-animated version for animated version, evaluations can be biased. For an example, see Kahneman (2011, p 119) for anchoring effects. The problem is that it is difficult to find fundamental objective anchor where UX evaluations can be based upon.

Regarding validity of AttrakDiff, an issue was found. AttrakDiff consists of semantic differentials where one adjective of the pair represents positive while other negative. However, it was found that adjective meant to represent negative UX is not always held as negative. One semantic differential item pair included in AttrakDiff measure was predictable–unpredictable, where predictable represents positive and unpredictable negative. One participant evaluated site as unpredictable, which lowered the summary variable of perceived UX. While evaluating, he explained that quality of being unpredictive was a positive thing for him, not bad. If there exists one that views matter in this light, it is possible that there exist others with similar views as well. This certainly reduces validity of AttrakDiff as a measurement tool of UX. It could be beneficial to study how people perceive semantic differential and modify AttrakDiff according to findings.

8 CONCLUSION

In here, conclusions of the study are summarised. Main findings of the study are restated and compared with prior literature. Additionally, limitations are presented and opportunities for future studies are listed.

8.1 Conclusions

The present study utilised user-centred theory of UX (Hassenzahl, 2008) and theory of experienced utility (Kahneman et al., 1997) for experimentally investigating how consistent users' retrospective evaluations are with hedonic real-time evaluations of website experience. Objective was to investigate how well memories correspond with actual subjective experience. Secondly, interest lied in examining explanatory power of sequencing effects to memories of experiential episode. Thirdly, we explored influence of animated transitions on UX. First two objectives were examined with within-subjects method and last one with between-subjects method. The study was initiated by will to contribute knowledge for practitioners regarding how good UX can be attained. Thesis also considered most important factors that determine UX.

The results of this study support Hassenzahl's (2008) view that emotions play pivotal role in formation of UX. Momentary feelings elicited during experiencing determine how experience with a product will be remembered, and they will serve as a central basis of product evaluations (Hassenzahl, 2008). However, some momentary feelings pertaining to experiential episode are more important than others. Kahneman et al. (1997) found that most intensive and ending moments are more impactful to memory than others. As an outcome, sometimes memories and decisions are not "correctly attuned to experience" (Kahneman, 2011, p. 384).

That outcome was not apparent in this study. The results indicated that users' memories of mostly positive experiences are not greatly biased. Although there were small differences, retrospective evaluations were generally consistent with real-time evaluations. It is beyond this paper to infer whether par-

ticipants formed retrospective evaluations by recollecting their memory of experience as was intention, by recollecting their memory of evaluations, or by some other means, such as the plausibility judgement strategy.

Results do not comply with prior finding that overall perceptions of experience do not reflect accurately average of emotions felt during experiencing (e.g., Kahneman et al., 1997; Kujala et al., 2011). In this study, hedonic real-time evaluations, on average, strongly corresponded with retrospective evaluations. Regardless, results did not contradict peak-end rule, as the most intense and last evaluation predicted, to a high degree, retrospective evaluation of an experience. Additionally, strong recency effect was found. Experimental design created conditions where, successive evaluations were consistent and peak-end effect was significant. Perhaps, participants did mostly evaluate past experience based on peak and end moments, but this evaluation was coincidentally consistent with whole actual experience. As participants mainly perceived website UX as positive, this coincidence is certainly plausible. Additionally, while peak-end effect was strong right after interaction, our results imply that the effect decreases by the passage of time.

Kraft and Hurtienne (2017) did not find significant effects of animated transitions to UX with within-subjects study design. Similarly, significant effects were not found in present study neither. Differences between website versions were trivial in both measurement points. Results suggest that in a quest to find central aspects of website UX, animated transitions can be left out of considerations. Yet, it is not inconceivable that positive effects, or even competitive edge can be attained by the means of animating layout transitions. More scholars should pursue this topic in order to scientifically determine effects of animated transitions.

This study integrated knowledge from various disciplines, such as HCI, cognitive psychology and neuroscience for contributing to understanding of UX. What can be learnt from sequencing effects and animated transitions? They allow us to understand that designing interactive products does not need to merely involve careful crafting of interactive elements, it can also include planning of temporal progression of UX. Like screenwriter writes a script for the movie, a UX designer can write a script for the interaction (Hassenzahl, 2010, p. 70). A good designer knows which moments of the interaction primarily contribute to the memory of the experiencer. As movies, interactions have beginning, end and peak moments. An interactive episode can be a journey, which contains drama arch where tension is established, built up, and eventually released. As the icing on a cake, the journey can be guided with a help from animated transitions which connect separate parts to a single coherent story. As a result, user's needs are met exquisitely. To conclude, by paying attention to temporal details, a designer can potentially elevate the experience to another level.

8.2 Limitations

The conduct of research was done with utmost care and attention for details. At any rate, that was the intention. However, as readers might have noted, there are severe limitations to present study. Above all, two limitations emerge before others. Firstly, the greatest limitation was insufficient sample size (N = 25). Author did not have adequately resources for gaining enough participants for the empirical study. Hence, one cannot make future predictions based on the data acquired from participants. It proved difficult to find participants for the study: recruiting participants via university's email list provided merely three participants. Best method of acquiring them proved to be asking people in person.

In retrospect, more effort could have put towards recruiting participants. Though, it is much easier to gain adequate sample size when data is collected only via online survey. Experiments, on the other hand, demand much more time and effort from both researcher and participants. There is need to find time and place that suits both parties. With online surveys, participant only needs to have internet connection for providing data for the researcher. Moreover, as in this master thesis UX over time was under investigation, participants had to fill another survey in addition to one they filled by the experiment.

The second prominent limitation, partly related to first one, occurred due to excessive ambition in including two fairly complex research topics in one research. If only consistency and sequencing effects were studied, there would have been more time to hone experimental design for investigating memory effects. If only animated transitions were studied, study could have been conducted with within-subjects design and we could have received more interesting results. Also, in that case sample size would not have been that big of an issue.

Since the time for designing the experiment was limited, and by its nature, laillisetpalvelut.fi did not seem to be very emotionally evocative website, there was a fear that the stimuli for momentary experience will be too weak for studying successive evaluations and sequencing effects. Weak momentary stimuli would cause everyone to monotonously rate zeros for every sub-interaction. Fortunately, that fear was found unfounded. However, problem was not that the website did not evoke any feelings — many participants rated maximum positive values for sub-interactions — but that all evaluations of single temporal profile did not vary that much. If participant felt that one interaction was pleasant, it was likely that other five interactions were pleasant as well. When evaluations do not vary from each other, there is not much room for occurrence nor examination of sequencing effects.

In second measurement point, participants might have actually tried to remember what they evaluated in first measurement point instead of trying to re-evaluate the same experience, as contemplated in 6.2. Moreover, addressing measurement of real-time evaluations, there might have been individual differences regarding how participants assumed and used pleasant-unpleasant scale. For instance, an intensity of emotion which represented number three for one

person could have been number one for other. Some people might give extreme evaluations more loosely than others.

For investigating effects of animated transitions, within-subjects design would have been more suitable than between-subjects used in present thesis. Within-subjects design was not chosen as it would have decreased validity of research questions 1 and 2. It was thought that participants cannot validly evaluate two different yet similar experiences after time period of one week. It could have been hard to differentiate memory of one experience from other.

With between-subjects design, it is challenging to measure differences between two versions of website. There can be countless of unnoticeable variables in the background that can affect the participant evaluations, other than differences between transitions. To get more valid results, sample size should be greater with between-subjects design.

Implementing animations to website proved to be surprisingly difficult and time-consuming. It required a great deal of effort and time to program transitions as intended for the needs of this study. Because of implementation took so much time, the whole research project left behind in schedule. Furthermore, due the difficulty of programming transitions, author was not able to implement precisely all the transitions as planned.

8.3 Future research

High correlations of real-time evaluations to retrospective evaluations are no surprise when most of the evaluations are positive. Hence, future studies should utilise more controversial websites or other study objects in experiments. At least, experiments should be designed in way that generates both positive and negative real-time evaluations during one experiential episode. Additionally, researchers should use websites that generate more variance among evaluations of participants, so that there would be both positive and negative evaluations. Unfortunately, it can be very difficult to predict which websites likely result in diverse evaluations.

As mentioned, both research topics, namely sequencing effects and animated transitions demand their own separate studies. Both topics are complicated enough by themselves. Therefore, future studies should choose to investigate only one of the two topics. What is certain, there remains research gaps in both domains that needs to be addressed in future. We cannot yet say final word about how consistent different UX memories tend to be, nor which sequencing effect is the most important. We do not have yet robust proof regarding how animated transitions influence UX.

There is a need for longitudinal studies examining development of UX (Hassenzahl, 2010, p. 21). If memory is consistent with experiencing after one week, does the consistency remain at the same level even after one month? What is left from the experience in memory after one year? Do there still remain some topmost feeling, as in Buehner's quote? There exist longitudinal studies where subjects interact more than once with the product (e.g., Karapanos et al.,

2009; Karapanos et al., 2010). It would be also valuable to understand how memory of a single experience progresses or deteriorate in one year.

Effects of animation should be investigated with within-subjects study design, as mentioned before. Both animated and non-animated versions of website should be shown to each participant, so that they could evaluate which version they prefer. Tasks included in interactions should be identical with both versions, and tasks should be simple and include sufficient number of layout transitions from one view to another. This method would give more insight into preferences of layout transitions.

Results of this study indicated that gender might be an influencing factor in perceptions of animated transitions. Future studies could examine more properly whether effect of animated transitions depends to any degree on gender of the user. Tuch, Bargas-Avila and Opwis (2010) have found gender differences in preferences of aesthetics of website design. Therefore, it would not be impossible that animated transitions, aesthetic elements themselves, are perceived differently between men and women.

Addressing co-experience approach and expectations, studying effect of positive or negative priming to UX evaluations, in a similar vein with Raita and Oulasvirta (2010), could be illustrative. How are evaluations affected by the different claims told to participants before interaction? This could be examined, for instance, by dividing participants to different groups, as between-subjects design. Participants of positive priming group would be told something along the lines of website regarded as state of the art and is award winner. Participants of negative priming group would be told that site was done by amateurs and is considered as distasteful and unusable. Participants belonging to third group would comprise control group and nothing would be told to them. Importantly, messages should be delivered convincingly to both priming groups.

Hassenzahl (2008) states that UX consists of pragmatic and hedonic aspects. Future studies could go beyond overall UX and investigate how perceptions of hedonic and pragmatic quality of experience develops over time. Moreover, it would be interesting to discover how animated transitions influence different qualities of UX. Which quality is more affected by the transitions, pragmatic or hedonic?

REFERENCES

Anderson, J. R. (2015). *Cognitive psychology and its implications*. Macmillan.

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

Baecker, R., & Small, I. (1990). Animation at the interface. *The art of human-computer interface design*, 251-267.

Bailey, B. P., & Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in human behavior*, *22*(4), 685-708.

Bargas-Avila, J. A., & Hornbæk, K. (2011, May). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2689-2698). ACM.

Battarbee, K., & Koskinen, I. (2005). Co-experience: user experience as interaction. *CoDesign*, *1*(1), 5-18.

Baumgartner, H., Sujan, M., & Padgett, D. (1997). Patterns of affective reactions to advertisements: The integration of moment-to-moment responses into overall judgments. *Journal of Marketing Research*, 219-232.

Chang, B. W., & Ungar, D. (1995). Animation: from cartoons to the user interface.

Bederson, B. B., & Boltman, A. (1999). Does animation help users build mental maps of spatial information?. In *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis' 99)* (pp. 28-35). IEEE.

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: a review/(Effet de l'animation sur les performances des utilisateurs: une sythèse). *Le travail humain*, *63*(4), 311.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276.

Brügger, N. (2009). Website history and the website as an object of study. *New Media & Society*, *11*(1-2), 115-132.

Brügger, N., & Schroeder, R. (Eds.). (2017). *The web as history: Using web archives to understand the past and the present*. UCL Press.

Carr, L. T. (1994). The strengths and weaknesses of quantitative and qualitative research: what method for nursing?. *Journal of advanced nursing*, *20*(4), 716-721.

Carver, C.S., Scheier, M.F., 1989. On the Self-Regulation of Behavior. Cambridge University Press, New York.

Chaffey, D., & Ellis-Chadwick, F. (2012). *Digital marketing: strategy, implementation and practice*(Vol. 5). Harlow: Pearson.

Chang, B. W., & Ungar, D. (1995). Animation: from cartoons to the user interface.

Chevalier, F., Dragicevic, P., & Franconeri, S. (2014). The not-so-staggering effect of staggered animated transitions on visual tracking. *IEEE transactions on visualization and computer graphics*, *20*(12), 2241-2250.

Chevalier, F., Riche, N. H., Plaisant, C., Chalbi, A., & Hurter, C. (2016, June). Animations 25 years later: new roles and opportunities. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (pp. 280-287). ACM.

Cockburn, A., Quinn, P., & Gutwin, C. (2015, April). Examining the peak-end effects of subjective experience. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 357-366). ACM.

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

Crafton, D. (2011). The veiled genealogies of animation and cinema. *Animation*, *6*(2), 93-110.

Czerwinski, M., Cutrell, E., & Horvitz, E. (2000, September). Instant messaging: Effects of relevance and timing. In *People and computers XIV: Proceedings of HCI* (Vol. 2, pp. 71-76).

Darwin, C., & Wallace, A. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the proceedings of the Linnean Society of London. Zoology*, *3*(9), 45-62.

Desmet, P. M., & Hekkert, P. (2007). Framework of product experience. *International journal of design*, *1*(1), 57-66.

Dessart, C. E., Genaro Motti, V., & Vanderdonckt, J. (2011, June). Showing user interface adaptivity by animated transitions. In *Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems* (pp. 95-104). ACM

Diener, E., Wirtz, D., & Oishi, S. (2001). End effects of rated life quality: The James Dean effect. *Psychological science*, *12*(2), 124-128.

Djajadiningrat, J. P., Gaver, W. W., & Fres, J. W. (2000). Interaction relabelling and extreme characters: methods for exploring aesthetic interactions. *In Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 66-71). ACM.

Do, A. M., Rupert, A. V., & Wolford, G. (2008). Evaluations of pleasurable experiences: The peak-end rule. *Psychonomic Bulletin & Review*, *15*(1), 96-98.

Dragicevic, P., Bezerianos, A., Javed, W., Elmqvist, N., & Fekete, J. D. (2011, May). Temporal distortion for animated transitions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2009-2018). ACM.

Ekman, P. (1992a). Are there basic emotions?.

Ekman, P. (1992b). An argument for basic emotions. *Cognition & emotion*, *6*(3-4), 169-200.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion review*, *3*(4), 364-370.

Forlizzi, J., & Battarbee, K. (2004, August). Understanding experience in interactive systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 261-268). ACM.

Forlizzi, J., & Ford, S. (2000, August). The building blocks of experience: an early framework for interaction designers. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*(pp. 419-423). ACM.

Garrett, J. J. (2010). *Elements of user experience, the: user-centered design for the web and beyond*. Pearson Education.

Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological research*, *50*(4), 243-250.

Gutwin, C., Rooke, C., Cockburn, A., Mandryk, R. L., & Lafreniere, B. (2016, May). Peak-end effects on player experience in casual games.

In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 5608-5619). ACM.

Guzmán-Vélez, E., Feinstein, J. S., & Tranel, D. (2014). Feelings without memory in Alzheimer disease. *Cognitive and behavioral neurology*, *27*(3), 117.

Harari, Y. N. (2014). *Sapiens: A brief history of humankind*. Random House. (ehkä kannattaa löytää uskottavammat lähteet korvaamaan)

Harbich, S., & Hassenzahl, M. (2017). User Experience in the Work Domain: A Longitudinal Field Study. *Interacting with Computers*, *29*(3), 306-324.

Harrison, C., Amento, B., Kuznetsov, S., & Bell, R. (2007, October). Rethinking the progress bar. In *Proceedings of the 20th annual ACM symposium on User interface software and technology* (pp. 115-118). ACM.

Harrison, C., Yeo, Z., & Hudson, S. E. (2010, April). Faster progress bars: manipulating perceived duration with visual augmentations. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1545-1548). ACM.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & information technology*, *25*(2), 91-97.

Hassenzahl, M. (2010). Experience design: Technology for all the right reasons. *Synthesis lectures on human-centered informatics*, *3*(1), 1-95.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-computer interaction*, *19*(4), 319-349.

Hassenzahl, M., & Monk, A. (2010). The inference of perceived usability from beauty. Human–Computer Interaction, 25(3), 235-260.

Hassenzahl, M. (2008, September). User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th Conference on l'Interaction Homme-Machine* (pp. 11-15). ACM.

Hassenzahl, M., & Sandweg, N. (2004, April). From mental effort to perceived usability: transforming experiences into summary assessments. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1283-1286). ACM.

Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, and interactive products–Facets of user experience. *Interacting with computers*, *22*(5), 353-362.

Hassenzahl, M., Wiklund-Engblom, A., Bengs, A., Hägglund, S., & Diefenbach, S. (2015). Experience-oriented and product-oriented evaluation: psychological need fulfillment, positive affect, and product

perception. *International journal of human-computer interaction*, *31*(8), 530-544.

Heer, J., & Robertson, G. (2007). Animated transitions in statistical data graphics. *IEEE transactions on visualization and computer graphics*, *13*(6), 1240-1247.

Hong, W., Thong, J. Y., & Tam, K. Y. (2007). How do Web users respond to non-banner-ads animation? The effects of task type and user experience. *Journal of the American Society for Information Science and Technology*, *58*(10), 1467-1482.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies*, *64*(2), 79-102.

ISO DIS 9241-210:2008. Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems (formerly known as 13407). International Organization for Standardization (ISO). Switzerland.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations.

Jiang, Z., Wang, W., Tan, B. C., & Yu, J. (2016). The determinants and impacts of aesthetics in users' first interaction with websites. *Journal of Management Information Systems, 33*(1), 229-259.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, *14*(2), 201-211.

Jokinen, J. P. (2015). User psychology of emotional user experience. *Jyväskylä studies in computing*, (213).

Jokinen, J. P., Silvennoinen, J., & Kujala, T. (2018). Relating Experience Goals With Visual User Interface Design. *Interacting with Computers*, *30*(5), 378-395.

Kahneman, D. (2011). Thinking, fast and slow. New York: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (2013). Choices, values, and frames. In *Handbook of the Fundamentals of Financial Decision Making: Part I*(pp. 269-278).

Kahneman, D. (2000). Experience utility and objective happiness: a moment-based approach. In: Kahneman, D., Tversky, A. (Eds.), Choices, Values and Frames. Cambridge University Press, New York .

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

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The quarterly journal of economics*, *112*(2), 375-406.

Karapanos, E. J. Zimmerman, J. Forlizzi, and J. Martens. (2009). User experience over time. In *In Proc. CHI '09, 729–738. ACM*.

Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J. B. (2010). Measuring the dynamics of remembered experience over time. *Interacting with Computers*, *22*(5), 328-335.

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion review*, *1*(2), 99-113.

Kraft, J. F., & Hurtienne, J. (2017, September). Transition animations support orientation in mobile interfaces without increased user effort. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (p. 17). ACM.

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., & Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. *Interacting with Computers*, *23*(5), 473-483.

Kujala, S., Vogel, M., Pohlmeyer, A. E., & Obrist, M. (2013, April). Lost in time: the meaning of temporal aspects in user experience. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 559-564). ACM.

Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, *43*, 35-48.Löwgren, J. (2009). Toward an articulation of interaction esthetics. *New Review of Hypermedia and Multimedia*, *15*(2), 129-146.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339.

Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International journal of human-computer studies*, *60*(3), 269-298.

Law, E. L., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: A survey approach.

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 719-728.

Law, E. L. C., van Schaik, P., & Roto, V. (2014). Attitudes towards user experience (UX) measurement. *International Journal of Human-Computer Studies*, *72*(6), 526-541.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science*, *9*.

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology, 25*(2), 115-126.

Maeda, J. (2006). *The laws of simplicity*. MIT press.

McCarthy, J., & Wright, P. (2004). Technology as experience. *interactions*, *11*(5), 42-43.

Mekler, E. D., & Hornbæk, K. (2016, May). Momentary pleasure or lasting meaning?: Distinguishing eudaimonic and hedonic user experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4509-4520). ACM.

Merz, B., Tuch, A. N., & Opwis, K. (2016, May). Perceived user experience of animated transitions in mobile user interfaces. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 3152-3158). ACM.

Miyata, Y., & Norman, D. A. (1986). The control of multiple activities. In D. A. Norman & S. W. Draper (Eds.), User centered system design: New perspectives on human–computer interaction. Hillsdale, NJ: Lawrence Erlbaum Associates

Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies, 68*(10), 689-709.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic Civitas Books.

Norman, D. A. (2009). THE WAY I SEE IT Memory is more important than actuality. *Interactions*, *16*(2), 24-26.

Norman, D. (2010). Why design education must change. *core77*, *11*, 26.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Constellation.

Norman, D., & Nielsen, J. (n.d.). The Definition of User Experience (UX). Retrieved from https://www.nngroup.com/articles/definition-user-experience/

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of experimental social psychology*, *45*(4), 867-872.

den Ouden, E., Yuan, L., Sonnemans, P. J., & Brombacher, A. C. (2006). Quality and reliability problems from a consumer's perspective: An increasing problem overlooked by businesses?. *Quality and Reliability Engineering International*, *22*(7), 821-838.

Partala, T., & Kallinen, A. (2011). Understanding the most satisfying and unsatisfying user experiences: Emotions, psychological needs, and context. *Interacting with computers*, *24*(1), 25-34.

Phelps, E. A., & Sharot, T. (2008). How (and why) emotion enhances the subjective sense of recollection. Current Directions in Psychological Science, 17(2), 147-152.

Plutchik, R., & Kellerman, H. (Eds.). (2013). *Biological foundations of emotion* (Vol. 3). Academic press.

Prinz, J. (2005). Are emotions feelings?. *Journal of Consciousness Studies*, *12*(8-9), 9-25.

Raita, E., & Oulasvirta, A. (2010, September). Too good to be bad: The effect of favorable expectations on usability perceptions. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, No. 26, pp. 2206-2210). Sage CA: Los Angeles, CA: SAGE Publications.

Redelmeier, D. A., & Kahneman, D. (1996a). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *66*(1), 3-8.

Redelmeier, D., Kahneman, D., & Kahneman, D. (1996b). Improving the Memory of a Colonoscopy. *University of California-Berkeley, working paper*.

Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, *89*(3), 250.

Reinecke, K., & Bernstein, A. (2011). Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(2), 8.

Roto, V. (2007). User experience from product creation perspective. *Towards a UX Manifesto*, *31*.

Rousi, R., & Silvennoinen, J. (2018). Simplicity and the art of something more: A cognitive-semiotic approach to simplicity and complexity in human-technology interaction and design experience. *Human technology*, *14*(1).

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., ... & Nummenmaa, L. (2015). Discrete neural signatures of basic emotions. *Cerebral cortex*, *26*(6), 2563-2573.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and emotion*, *23*(7), 1307-1351.

Seligman, M. E., & Csikszentmihalyi, M. (2000). *Positive psychology: An introduction* (Vol. 55, No. 1, p. 5). American Psychological Association.

Sheldon, K. M., Elliot, A. J., Kim, Y., & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of personality and social psychology*, *80*(2), 325.

Silvennoinen, J. (2017). Apperceiving visual elements in human-technology interaction design. *Jyväskylä studies in computing*, (261).

Smith, C. A., & Lazarus, R. S. (1993). Appraisal components, core relational themes, and the emotions. *Cognition & Emotion*, *7*(3-4), 233-269.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of personality and social psychology*, *48*(4), 813.

Speier, C., Valacich, J. S., & Vessey, I. (1999). The influence of task interruption on individual decision making: an information overload perspective. Decision Sciences, 30(2), 337–360.

Sward, D., & Macarthur, G. (2007, September). Making user experience a business strategy. In *E. Law et al.(eds.), Proceedings of the Workshop on Towards a UX Manifesto* (Vol. 3, pp. 35-40).

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

Tambini, A., Rimmele, U., Phelps, E. A., & Davachi, L. (2017). Emotional brain states carry over and enhance future memory formation. *Nature neuroscience*, *20*(2), 271.

Thielsch, M. T., & Hirschfeld, G. (2012). Spatial frequencies in aesthetic website evaluations–explaining how ultra-rapid evaluations are formed. *Ergonomics*, *55*(7), 731-742.

Thomas, B. H., & Calder, P. (2001). Applying cartoon animation techniques to graphical user interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *8*(3), 198-222.

Thomas, F., & Johnston, O. (1981). *Disney animation: The illusion of life*. Abbeville Press.

Tomasello, M. (2014). The ultra-social animal. *European journal of social psychology*, *44*(3), 187-194.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13*(2), 127-145.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate?. *International journal of human-computer studies*, *57*(4), 247-262.

Tuch, A. N., Bargas-Avila, J. A., & Opwis, K. (2010). Symmetry and aesthetics in website design: It'sa man's business. *Computers in Human Behavior*, *26*(6), 1831-1837.

Tuch, A. N., Presslaber, E. E., StöCklin, M., Opwis, K., & Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies*, *70*(11), 794-811.

UEQ Questionnaire, 2017. Retrieved May 15, 2017 from http://www.ueq-online.org.

APPENDIX 1 INSTRUCTIONS FOR THE PROCEDURE (IN FINNISH)

1. Etusivuun Tutustuminen (ensivaikutelma)
    a. Tutustu rauhassa etusivuun ja tutki sen yleisilmettä
    b. Mieti mikä mahtaisi olla tämän palvelun tarkoitus. Välittyykö tämä tieto etusivulta.
2. Lisätietoa-sivut
    a. Mene *tietoja meistä* -sivulle, ja lue pikaisesti kuvaus sivustosta.
    b. Pohdi vastasiko kuvaus etusivulla saamaasi kuvaa sivuston luonteesta?
    c. Mene *kuinka tunnistaa piraattipalvelu* -sivulle, ja ota selvää laittomien ja laillisten palveluiden tuntomerkeistä
3. Ajankohtaista osio
    a. Palaa takaisin etusivulle ja etsi ajankohtaista-osio, mistä löytyy Facebook-päivityksiä. (älä klikkaa linkeistä)
    b. Selaa ja käy läpi kaikki kuusi viimeisintä päivitystä, jotka ovat sivulle listattu.
    c. Pohdi mielessäsi, mikä on kaikista mielenkiintoisin päivitys.
4. Haku.
    a. Pysy etusivulla
    b. Kokeile etsiä ensimmäistä mieleesi tulevaa verkkopalvelua.
        i. Löytyikö palvelu?
        ii. Jos ei löydy, etsi joku toinen alasvetovalikossa listattu palvelu ja klikkaa siitä.
    c. Tällä kertaa pysyen selaa palveluita -sivulla, etsi palvelua 'eMusic'
        i. Tutki mitä tietoja palvelusta annetaan.
    d. Selaa palveluita -sivulla, kokeile vielä etsiä palvelua nimeltä 'Hulu'
5. Filtterit. Part 1.
    a. Pysy selaa-palveluita-sivulla.
    b. Tutki, minkä kategorian (musiikki, e-kirjat, jne) palveluita löytyy eniten ja minkä vähiten?
    c. Tutki löytyykö sellaista palvelua, joka mahdollistaa musiikin kuuntelun, radion kuuntelun sekä elokuvien katselun? Miten tällaisen voisi löytää helpoiten?
6. Filtterit. Part 2. (viimeiset hetket)
    a. Pysy selaa-palveluita-sivulla.
    b. Etsi kaikki suoratoistopalvelut, joiden avulla voit kuunnella musiikkia tai katsella elokuvia ja tv-sarjoja.
        i. Kuinka monta palvelua löytyi?
    c. Etsi kaikki palvelut, joiden avulla voit pelata pelejä ilmaiseksi.
        i. Kuinka monta palvelua löytyi?

Etsi löytyykö sellaista suoratoistopalvelua, joka mahdollistaa e-kirjojen lukemisen tilauksesta.

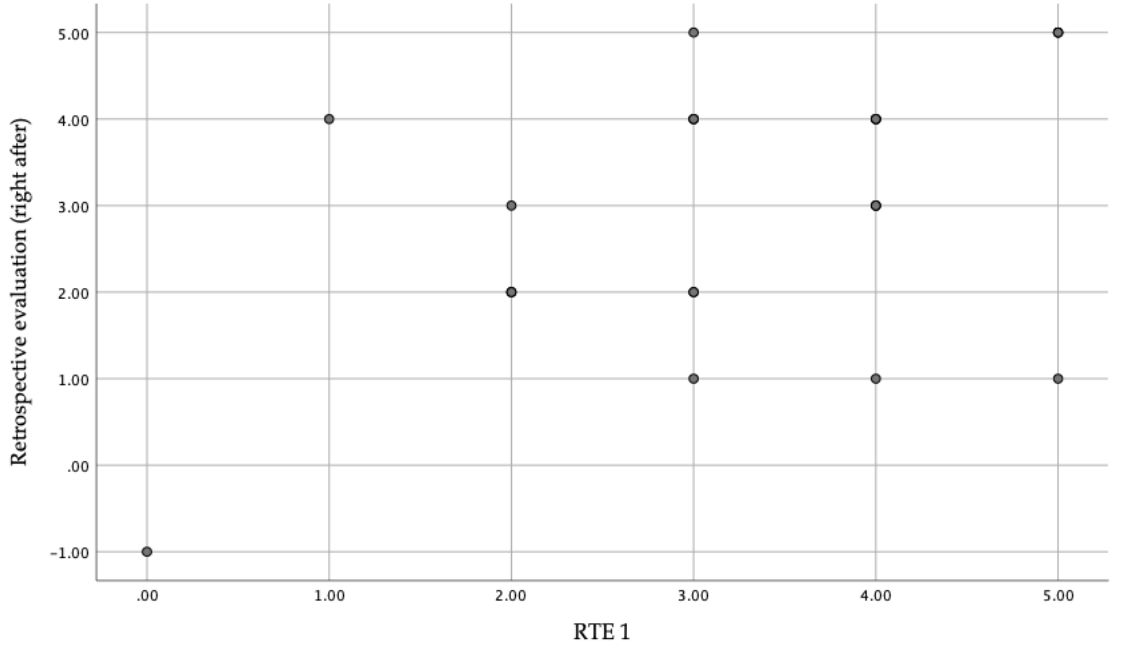
APPENDIX 2 ATTRAKDIFF 2 LITE

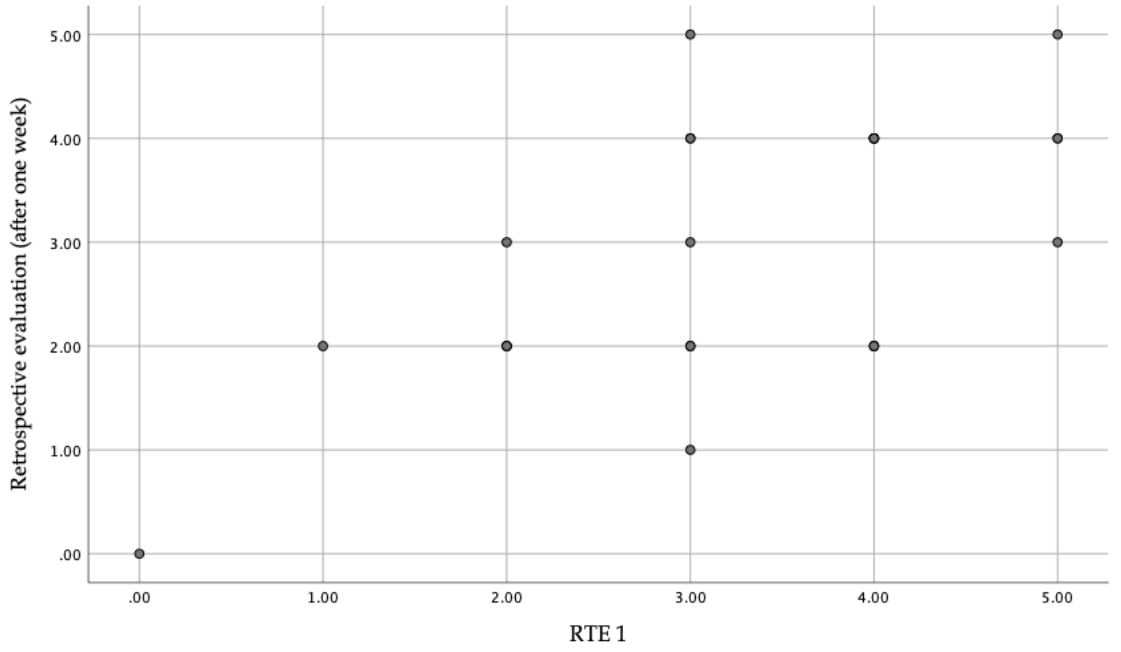
From Hassenzahl, M., & Monk, A. (2010). The inference of perceived usability from beauty. Human–Computer Interaction, 25(3), 235-260.

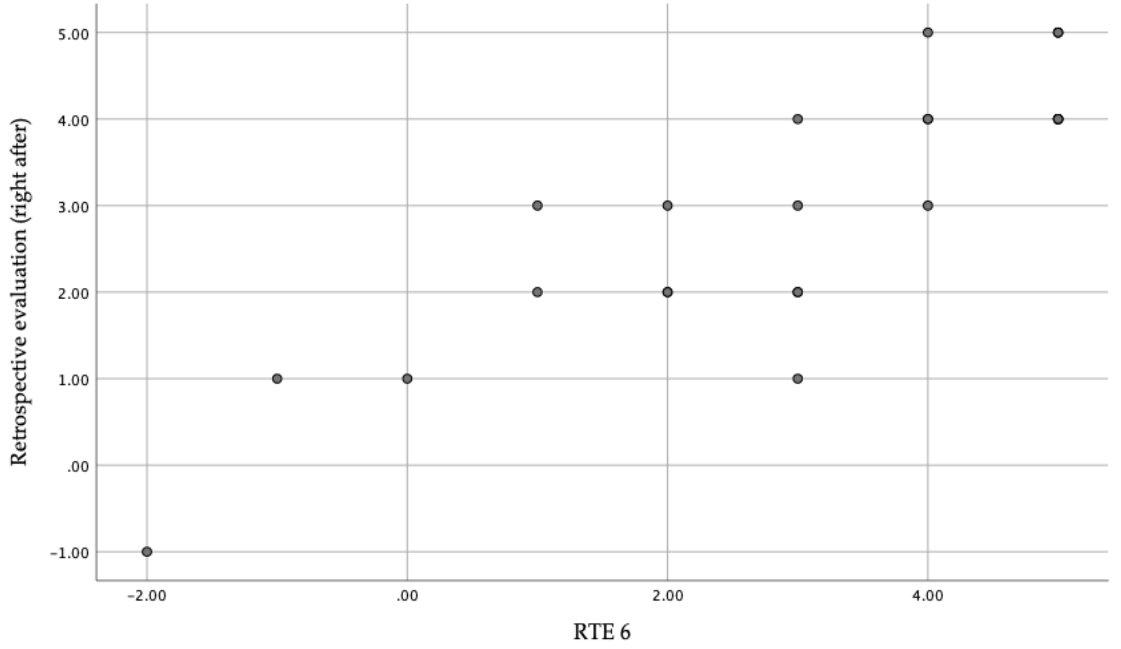
1. GOODNESS bad – good
2. BEAUTY ugly – beautiful
3. PQ 1 confusing – structured
4. PQ 2[a] impractical – practical
5. PQ 3 unpredictable – predictable
6. PQ 4 complicated – simple
7. HQ 1 dull – captivating
8. HQ 2[a] tacky – stylish
9. HQ 3 cheap – premium
10. HQ 4[a] unimaginative – creative

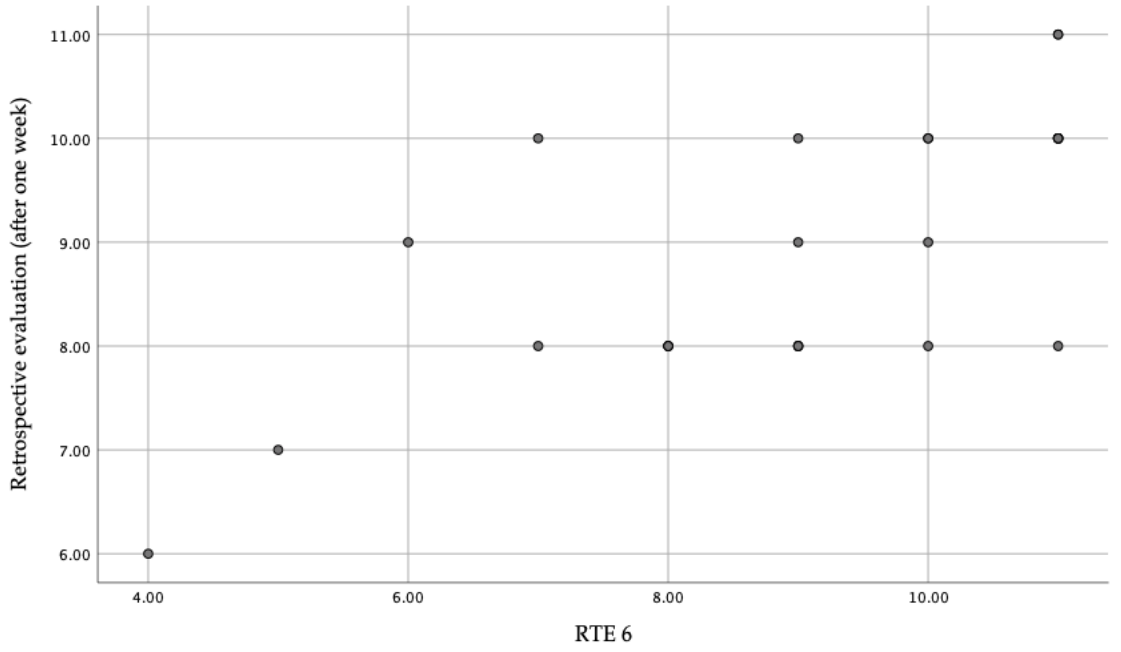
[a]Item is reversed.
PQ = pragmatic quality
HQ = hedonic quality

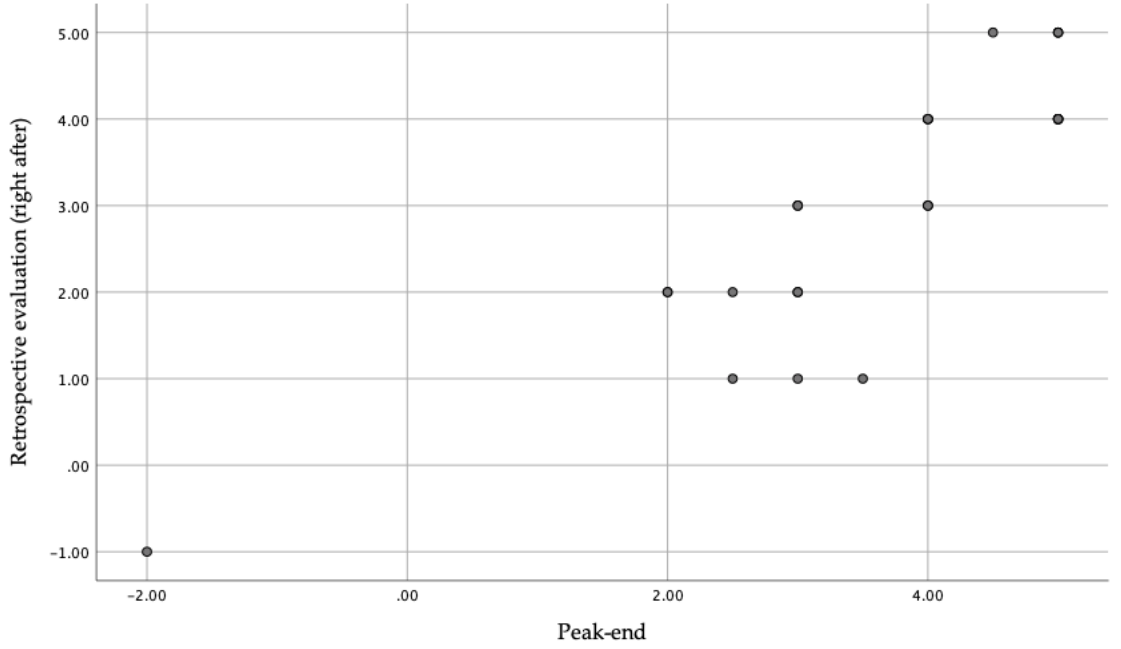
APPENDIX 3 SCATTERPLOT DIAGRAMS

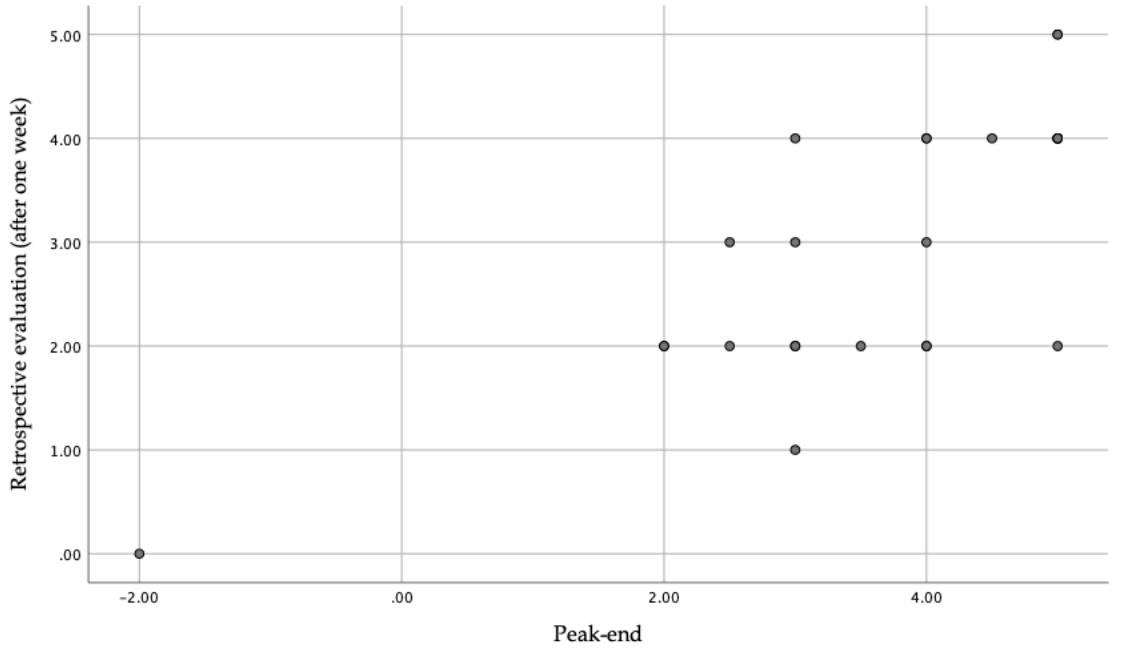





Retrospective evaluation (right after)
5.00
4.00
3.00
2.00
1.00
.00
-1.00
-2.00
.00
2.00
4.00
RTE 6


Retrospective evaluation (after one week)
11.00
10.00
9.00
8.00
7.00
6.00
4.00
6.00
8.00
10.00
RTE 6


Retrospective evaluation (right after)
5.00
4.00
3.00
2.00
1.00
.00
-1.00
-2.00
.00
2.00
4.00
Peak-end


Retrospective evaluation (after one week)
5.00
4.00
3.00
2.00
1.00
.00
-2.00
.00
2.00
4.00
Peak-end