Titre: Title:	Design, Development, and Validation of Social Scales for the UX Evaluation of Interactive Products
Auteur: Author:	Ehsan Mortazavi
Date:	2022
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Mortazavi, E. (2022). Design, Development, and Validation of Social Scales for the UX Evaluation of Interactive Products [Ph.D. thesis, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/10689/

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Design, development, and validation of social scales for the UX evaluation of interactive products

EHSAN MORTAZAVI

Département de mathématiques et de génie industriel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ doctor*

Génie industriel

November 2022

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

Design, development, and validation of social scales for the UX evaluation of interactive products

présentée par **Ehsan MORTAZAVI**

en vue de l'obtention du diplôme de Philosophiæ Doctor

a été dûment acceptée par le jury d'examen constitué de :

Jinghui CHENG, PhD, président
Daniel IMBEAU, PhD, membre et directeur de recherche
Jean-Marc ROBERT, Doctorat, membre et codirecteur de recherche
Philippe DOYON-POULIN, PhD, membre et codirecteur de recherche
Michel C. DESMARAIS, PhD, membre
Christian BASTIEN, Doctorat, membre externe

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my sincere gratitude to my supervisors, Dr. Daniel Imbeau, Dr. Jean-Marc Robert, and Dr. Philippe Doyon-Poulin. Their guidance, support, and advice helped me to finish this journey. I am thankful to Dr. Imbeau for giving me the opportunity to start my PhD at Polytechnique; to Dr. Robert for introducing me to the amazing field of user experience and his continuing support; and to Dr. Doyon-Poulin for believing in me and encouraging me, for without him my PhD could not have been accomplished.

I would like to thank the president and members of the jury for taking the time to evaluate this thesis. I am also thankful to all the users and experts who participated in the different data collections for this work.

Life is full of ups and downs, and I prefer not to forget that. That is why I would like to express my appreciation to my psychologist, Mrs. Razzaghi, who saved me from the dark days of depression.

My heartfelt appreciation goes to my parents, without their unwavering dedication and unconditional love I could not be here. Although we are geographically apart, you are always in my heart.

Above all, I am thankful to my wife, Mitra, who has been by my side through every step of this journey. She is a loving partner, a true friend, and a strong supporter who helped me to believe in myself and stayed strong when we faced different and unexpected challenges. I am grateful to have you in my life.

This research was funded by NSERC (Natural Sciences and Engineering Research Council of Canada) Individual Research Grant awarded to Dr. Jean-Marc Robert (#RGPIN-2018-06733).

RÉSUMÉ

L'expérience utilisateur (dont l'abréviation anglaise est UX) englobe tous les aspects de l'interaction entre un utilisateur et un produit, service ou système (seul le terme produit sera utilisé ci-dessous). Cela va du moment où une personne découvre un nouveau produit pour la première fois jusqu'à son utilisation et inclut ses réflexions sur ses interactions avec le produit. Au fil des années, l'UX est devenue incontournable pour le succès des produits commerciaux dans un marché concurrentiel. Pouvoir évaluer l'UX globale d'un utilisateur avec un produit ainsi que des dimensions spécifiques –de celle-ci est essentiel pour améliorer la conception du produit. Dans cette thèse composée de trois articles, nous avons fait le point sur l'état de l'art des outils d'évaluation subjective de l'UX et nous nous sommes concentrés sur la dimension sociale de l'UX qui est négligée dans les outils actuels. Nous avons mis en évidence l'importance de cette dimension pour différentes catégories de produits et développé quatre échelles pour l'évaluer.

Dans le premier article, nous avons effectué une revue systématique de la littérature (RSL) et examiné 325 études UX publiées au cours des 11 dernières années. Nos principaux résultats ont montré que 104 outils différents sont disponibles pour l'évaluation subjective UX, qu'ils peuvent être classés comme généraux ou spécifiques à un domaine, qu'ils sont applicables à une grande variété de produits et qu'ils couvrent plus de 300 dimensions UX. Notre catégorisation des dimensions UX en 13 dimensions principales a montré que les dimensions informationnelle, sociale, physique et cognitive étaient moins fréquemment présentes dans les outils actuels. Nous avons également constaté que les outils modulaires sont plus appréciés que les outils généraux d'évaluation UX. Parce qu'ils sont plus complets, flexibles, faciles à utiliser, peu coûteux, rapides et extensibles. Par conséquent, après avoir identifié les dimensions UX manquantes et la meilleure approche d'évaluation, nous avons décidé de développer un module formé d'échelles sociales d'évaluation.

Bien que les outils modulaires facilitent l'évaluation, comment choisir les dimensions UX à évaluer nécessitait une enquête plus approfondie. Dans le deuxième article, nous avons étudié quelles dimensions UX sont plus importantes pour l'évaluation de différentes catégories de produits logiciels. Afin de savoir si le type d'utilisateurs évaluant l'UX et la culture ont un impact sur la perception de l'importance d'une dimension UX, nous avons demandé à 200 utilisateurs finaux et

huit experts UX du Canada d'évaluer 21 dimensions UX, incluant la dimension sociale, pour 15 catégories de produits. L'étude a montré que les résultats des évaluations de l'importance des dimensions UX par les utilisateurs finaux étaient similaires à ceux des experts dans 77 % des cas, alors qu'au sein du groupe des utilisateurs finaux, leurs résultats étaient similaires dans 97 % des cas. Nous avons également comparé nos résultats à ceux d'autres pays et confirmé que la perception de l'importance des dimensions UX était déterminée par la catégorie de produits, et non par la culture des répondants.

Trouver la relation entre l'identité, la sociabilité et l'acceptation sociale, et déterminer leur importance pour certaines catégories de produits ont été à la base du troisième article, qui portait sur le développement des échelles sociales.

Nous avons analysé la dimension sociale qui était présente ou absente dans les outils d'évaluation subjective UX disponibles et ce faisant, les sous-dimensions qui en faisaient partie. Nous avons discuté les résultats de cette analyse avec sept experts UX et avons ainsi dressé une liste de 27 items pour l'évaluation de la dimension sociale. Au moyen d'un questionnaire, 229 participants ont évalué leur UX avec des produits ayant des aspects sociaux; nous avons identifié quatre échelles sociales avec un total de 16 items à la suite d'une analyse factorielle exploratoire, à savoir l'identification, l'interaction sociale, la stimulation sociale et l'acceptation sociale. Ensuite, nous avons réalisé une étude de validation avec 450 participants, évaluant l'UX relative à trois différentes catégories de produits allant de très social à légèrement social. Les échelles étaient fiables avec un score alpha de Cronbach compris entre 0,792 et 0,846. De plus, des comparaisons avec l'outil d'évaluation UX AttrakDiff et six dimensions UX de l'UEQ+ ont révélé qu'il y avait des corrélations positives entre les nouvelles échelles sociales et la dimension d'identification d'AttrakDiff (0,24 - 0,69), et la dimension d'attractivité d'AttrakDiff (0,18 - 0,66) et UEQ+ (0,21 - 0,67) ($p<0,05$).

Nous estimons que la principale contribution de cette thèse a été le développement et la validation de quatre échelles sociales qui peuvent être utilisées en combinaison avec d'autres échelles de l'outil modulaire UEQ+ pour évaluer la dimension sociale de l'UX avec différents produits.

ABSTRACT

User eXperience (UX) encompasses all aspects of the interaction between a user and a product, service, or system (only the term product will be used below). From the moment a person first learns of a new product, to using it, and later reflecting on the interactions with the product, all these events contribute to UX. Over the years, UX has grown to become unavoidable for the success of commercial products in a competitive market. Being able to evaluate the overall UX of a user with a product as well as its specific dimensions are keys to improve the design of a product. In this thesis composed of three articles, we reviewed the state of the art in UX subjective evaluation tools and focused on the social dimension of UX, which is overlooked in the current tools. We highlighted its importance for different product categories and developed four scales to evaluate it.

In the first article, we did a Systematic Literature Review (SLR) and investigated 325 UX studies published during the last 11 years. Our main results showed that 104 different tools are available for UX subjective evaluation, that they can be classified as general or domain-specific, that they are applicable for a wide variety of products and cover more than 300 UX dimensions. Our categorization of UX dimensions under 13 main dimensions showed that the informational, social, physical, and cognitive dimensions appeared to be less frequently present in current tools. We also found that modular tools have become appreciated among general UX evaluation tools. It is because they are more comprehensive, flexible, easy to use, low-cost, rapid, and extendable. Therefore, after identifying the lacking UX dimensions and the best evaluation approach, we decided to put our focus on developing modular social scales.

Although modular tools facilitate the evaluation, the question of which UX dimensions should be evaluated needed more investigation. In the second article, we studied which UX dimensions are important for the evaluation of different software product categories. In order to find whether the perception of the importance of a UX dimension is influenced by the type of user or culture, we asked 200 end users and eight UX experts from Canada to evaluate 21 UX dimensions, including the social dimension, for 15 product categories. Our results showed that the importance ratings of UX dimensions between end users and experts were similar in 77% of cases, whereas within end-users' their ratings were similar in 97% of cases. We also compared our results to those of other

countries and confirmed that the importance of UX dimensions is determined by the product category, not the culture of the respondents.

Finding the relationship between identity, sociability, and social acceptance, and their importance for some product categories were the foundation for the final article, which was the development of the social scales. In this article, we analyzed the available UX subjective evaluation tools regarding the social sub-dimensions they incorporated. We discussed our findings with seven UX experts and came up with a list of 27 items for the evaluation of the social dimension. Through a questionnaire, 229 participants rated their UX with products having social aspects, and we identified four social scales with a total of 16 items as the result of exploratory factor analysis, i.e., identification, social interaction, social stimulation, and social acceptance. Next, we did a validation study with 450 participants, evaluating the UX of three different product categories ranging from highly social to slightly social. The scales were reliable with Cronbach's alpha scoring between 0.792 and 0.846. Moreover, comparisons with AttrakDiff, and six UX dimensions of UEQ+ reported the positive correlations between the new social scales and identification dimension of AttrakDiff (0.24 - 0.69) as well as with the attractiveness dimensions of AttrakDiff (0.18 - 0.66) and UEQ+ (0.21 - 0.67) ($p<.05$).

In conclusion, we estimate that the main results of this thesis are four validated social scales that can be used in combination with other scales of the UEQ+ modular framework to evaluate the social dimension of UX with different products.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AMT	Amazon Mechanical Turk
AR	Augmented Reality
CFA	Confirmatory Factor Analysis
CUE-model	Components of User Experience model
EFA	Exploratory Factor Analysis
GEQ	Game Experience Questionnaire
HCI	Human-Computer Interaction
ISO	International Organization for Standardization
meCUE	modular evaluation of key Components of User Experience
PANAS	Positive and Negative Affect Schedule
SAM	Self-Assessment Manikin
SLR	Systematic Literature Review
UEQ	User Experience Questionnaire
UEQ+	User Experience Questionnaire plus
UX	User eXperience
VR	Virtual Reality

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

1.1 Context

From turning off the alarm in the morning to watching movies on a video streaming service at night, people are constantly exposed to different products[1], services, or systems throughout the day. As a result, every aspect of user-product interaction has become increasingly important. Users' expectations have also increased and gone beyond performance, ease of use, and efficiency to seek aesthetically pleasing products that are enjoyable to use and evoke positive emotions. These aspects are addressed by the user experience (UX), an umbrella term that incorporates all aspects of the interaction between a user and a product.

UX is comprehensive yet ambiguous in that it has been considered as something desirable without being clearly defined [1], something that cannot be designed but can be designed for [2]. As UX does not belong to one single domain, a variety of definitions have been proposed over the years. We adopted Robert and Lesage's definition of UX: "a multidimensional construct that defines the overall effect over time on the user of interacting with a system and service in a specific context". This definition emphasizes that UX is the result of interacting with a product, addressing the main criticism of the UX definition of ISO 9241 [3]. ISO defines UX as "a person's perceptions and responses that result from the use and/or anticipated use of a product, system, or service" [4].

The multidimensional characteristic of UX results in the use of different evaluation methods and corresponding tools. Behavioral, physiological, and subjective evaluation methods provide real-time, retrospective, qualitative, quantitative, and mixed data. In UX evaluation studies, subjective evaluation tools such as questionnaires and scales are most commonly used. These tools can help document the user's emotions, expectations, perceptions, judgments, satisfaction, and values. This research is focused on subjective evaluation tools, and the following chapters help understand UX by studying its dimensions and addressing the shortcomings of the current evaluation tools.

---

[1] The word products will be used hereafter to refer to products, services, and systems, except for the UX definition section.

1.2 Definition of the research problematic

It has been more than two decades since UX was introduced, and during this time, numerous empirical and theoretical studies on this subject have been published. This thesis is more concerned with UX evaluation tools and, in particular, the evaluation of specific UX dimensions. Indeed, UX is multidimensional and the importance of different dimensions to the overall experience varies between product categories. Usefulness, efficiency, and perspicuity are prime dimensions for a word processor; immersion and stimulation are sought after in video games, whereas social dimension is underpinning social network apps.

The multidimensional characteristic of UX resulted in the use of many evaluation tools, and the UX dimensions covered by these tools varied as well. There are some tools that measure only one dimension, like emotion, while others evaluate multiple dimensions. The number and range of products evaluated in recent years have increased [5], and the new technologies and products have introduced new UX dimensions. For instance, narrative immersion for interactive digital narratives; and response quality and comprehensibility for voice digital assistants are UX dimensions that have not been previously considered in UX evaluation tools. The sheer number of UX evaluation tools and UX dimensions calls for a comprehensive review of UX empirical studies to identify which products and fields of use are being studied, how effective UX evaluation tools are regarding the UX dimensions they incorporate, and the way they evaluate different products.

The analysis of UX evaluation tools reveals that most of them give equal importance (weight) to each dimension. The benefit of assigning a degree of importance to each UX dimension is twofold. First, it helps designers to prioritize the most important dimensions when developing a new version of a product, particularly when there is a limited time or budget. Second, it assists novice evaluators in selecting the most important UX dimensions when using modular evaluation tools. Therefore, providing a list of important UX dimensions for different product categories improves the design and evaluation processes.

AttrakDiff [6], UEQ [7], and meCUE [8] are well-established questionnaires for the UX evaluation of a variety of products [9]. It is imperative to examine whether the UX dimensions they cover are comprehensive in the midst of the rapid development of new products. For instance, analyzing these tools regarding their coverage of social dimension of UX shows that AttrakDiff and meCUE

only evaluate identification while UEQ — the most commonly used subjective evaluation tool — does not cover any social aspect. Considering that identification is only one sub-dimension of the social dimension indicates that even well-established tools have limitations that should be addressed.

1.3 Objectives

Considering the problems outlined above, this thesis aims to accomplish three objectives. The **first objective** is to provide the state of the art on UX evaluation tools and UX dimensions. Performing a Systematic Literature Review (SLR) in article 1, we identified the products, field of use, study type, evaluation time, duration of interaction, types of collected data, data collection methods, evaluation tools, and the UX dimensions that have been studied in recent empirical studies. We defined the other objectives based on two findings of the SLR. First, results showed that informational, social, physical, and cognitive dimensions are underrepresented in current UX evaluation tools. Second, we found modularity as the emerging characteristic among the generally applicable UX evaluation tools (i.e., meCUE, UEQ+). Therefore, we decided to develop social scales for the UEQ+ modular framework. This choice was made because social features are making their way into a growing number of interactive products and UEQ+, as an extendable modular tool, was a promising framework to include this dimension.

Modular evaluation tools burden the evaluators with the responsibility of selecting the correct modules for the evaluation of their products. The **second objective** of this thesis is to investigate the importance of UX dimensions for different product categories. In article 2, we included the social dimension among the 21 UX dimensions to be evaluated for 15 product categories by end users and UX experts. We aim to compare the ratings of end users and UX experts, to find the relationship of social dimension with other dimensions, to compare our results from Canada to those of two other countries, and to provide a list of important UX dimensions for each product category. Results can help evaluators who use modular evaluation tools like UEQ+ to have a better selection of UX dimensions.

The **third objective** is to develop and validate social scales for the evaluation of interactive products based on the UEQ+ framework. Considering the problems outlined above, the overall objective of this thesis is to improve the UX subjective evaluation tools for interactive products.

This thesis accomplished this goal with three sub-objectives. In the first article we showed the lack of attention to social dimension and the potential of the UEQ+ framework; In the second article we coupled this information with the importance rating of UX dimensions for different product categories and found the presence of three social sub-dimensions. Together, this built the foundation for developing and validating social scales in the third article. The sum of these contributions offered validated scales to evaluate the social dimension of UX with interactive products, hence fulfilling this thesis' overall objective.

CHAPTER 2 LITERATURE REVIEW

This chapter provides an overview of the UX literature. UX scope, definition, and characteristics sections describe what UX is and how it differs from other HCI concepts. The remaining sections discuss the UX models, dimensions, evaluation methodologies, and tools. Chapter 4 (article 1) delves deeper and provides the state of the art in UX by conducting a Systematic Literature Review (SLR) on UX dimensions and subjective UX evaluation tools.

2.1 UX scope

The term "user experience" was first coined by Donald Norman in the 1990s [10]. In his opinion, human interface and usability were not capable of capturing all aspects of a person's experience because they offered a narrow insight into human-computer interaction [11]. Over the last two decades, UX has advanced in the field of HCI [12] and has evolved from what was once considered a fad, fuzzy, fashion, and buzzword [1, 13] to a core concept of HCI [14] that has been used in numerous empirical and theoretical studies [15]. The variety of products and use-contexts that have been studied highlights the expansion of UX boundaries [16, 17].

It was Hassenzahl's [18] concept of hedonic quality that added an appealing and enjoyable aspect to interaction with computers that were, at the time, only professional and serious tools [19]. As studies on experience, emotion, pleasure, and beauty began to expand the scope of HCI, the spotlight moved away from usability and its task-related approach [19]. According to some researchers, UX is limited to interactions between a user and a product via a user interface [1, 12, 16]. Others, however, disagree and argue that the scope of UX applies to all the products of everyday life [20]. Besides including pragmatic and non-interactive products [19], the increasing number of studies evaluating prototypes and services shows that the boundaries of UX are expanding [5].

The changes in industries and users made the emergence of UX inevitable. On the one hand, technology advances and increased manufacturing capabilities enabled companies to offer products with high levels of usability and functionality, while the shift from mass-production to mass-customization [21] made attention to details a must in a competitive market. Furthermore, users' expectations have changed, and usability as a means to remove dissatisfaction [22] is insufficient

for users. Besides, people's higher living standards, the diversity and complexity of their needs, and the variety of experiences offered by new products have led to an increase in usability being taken for granted [23]. Jordan's hierarchy of quality is applicable here, which states that products should fulfill functionality and usability prior to hedonic aspects [24]. As a result, competition is now beyond usability.

The fact that UX has a wide coverage and does not belong to a single domain makes it a difficult concept to define, measure, and evaluate. It is possible to see UX at the intersection of psychology, design, and information technology [22]. However, the diverse backgrounds of UX professionals and how they linked UX to their disciplines increased the complexity. According to [25], HCI and cognitive science connect UX to computer science and psychology. Graphic artists and software designers work with UX in web design and interaction design. Branding professionals deal with UX under experience design and marketing researchers in user analytics. UX is also linked to information architecture, user research, and human factors through the library science, anthropology, and ergonomics fields [25].

UX is sometimes used as a synonym for interaction and user-centered design [16] or used interchangeably with usability [13]. It could be due to the fact that UX lacks a scientific consensus on a definition [26]. However, to better understand the relationship between UX and usability, three approaches can be observed [27]:

- *UX is a component of usability*: based on this approach, UX is the elaboration of the satisfaction attribute of usability [26] which is responsible for subjective qualities such as likability, pleasure, comfort, and trust [28].
- *UX includes usability*: This view considers UX as a holistic approach which covers usability under terms like pragmatic or instrumental [14].
- *UX complements usability*: This approach believes that UX is larger than the satisfaction and attitude dimensions of usability and has expanded usability by adding subjective, emotional, and temporal aspects [14].

UX and usability are inextricably linked [12]. In a sense, UX owes its evolution to different trends such as usability, distributed cognition, emotion design, and activity theory [14]. However, studies have also considered UX as a counter-movement to usability thinking by comparing their core

meanings in human-computer interaction, dealing with efficiency and hedonic aspects, respectively [29]. Thus, further elucidation of UX definitions, characteristics, dimensions, and evaluations can help differentiate UX from other concepts in HCI.

2.2 UX definition

The importance of defining a concept rises from the statement of Fenton and Pfleeger that says "you cannot control what you cannot measure and you cannot measure what you cannot define" [30]. In light of this, finding a standardized and commonly agreed-upon definition of UX would have the following benefits [24]:

- Facilitate the teaching and dissemination of the UX concept.
- Improve scientific communication between UX researchers from different disciplines.
- Provide better management and control of the UX [31].

Efforts have been made to reach a universal definition of UX [32]. However, the multidisciplinary nature of UX has led to different definitions [33]. The diversity of backgrounds, interests, and perspectives among researchers and practitioners rules out the possibility of obtaining a universal definition [1]. The wide range of UX dimensions and different units of analysis (e.g., single user, multiple users, single product, multiple systems) are the other contributing factors [1]. Despite the benefits of a standardized definition of UX, there is a concern that it could be useful for only a limited time because the fast changes of the internet age defy sticking to a fixed definition [31].

Experience cannot be precisely defined [34]; nevertheless, we can differentiate user experience from general experience. Walking in a park, talking to someone, or watching a sport event are not included in the UX scope because they focus on the broad sense of experience. However, the interaction between the user and the product in each situation is considered a user experience (i.e., the shoes used for walking, the phone used for talking, and the television used for watching the sport event).

In UX, there are two types of users (active and passive) and three types of systems (non-interactive, adjustable, and interactive) [35]. Active users are the target users of a system who directly interact with it. Passive users, however, receive a second-hand or expected UX as a result of exposure to

advertisements or direct observation of others using a product [35]. Non-interactive systems are often high-tech systems that do not require control on the user's part with a simple interaction scenario (e.g., a city bus). The adjustable systems are mechanical or digital systems that require one-time or constant adjustments over the course of the interaction (e.g., an office chair, an air conditioner). Finally, the interactive systems need three or more dialogues between the user and the system to be separated from the adjustable systems. The level of interactivity can be as high as video games or lower levels such as feature phones, or ATMs [35].

In addition to general experience, it is important to understand and distinguish UX from other types of experiences. The following is a description of four other types of experience:

- Brand experience "includes not only interaction with the branded products, but interaction with the company, its products and services". Brand experience is a broader concept than UX. Brand experience can be influenced by advertisements, reviews, and people's opinions. It affects the actual usage and the feelings of users towards a product. Similarly, after a real encounter, the UX of a product typically impacts the brand experience [1].
- Product experience refers to interaction with an artifact. By focusing only on commercial products, it has a narrower scope than UX. [1].
- Co-experience, shared experience, and group experience refer to "situations in which experiences are interpreted as being situated and socially constructed" [33]. The focus is placed on the interaction with the system and on the people who create and experience a situation together.
- Service experience "in a broad sense can refer to face-to-face services (e.g., in a restaurant or repair point), public services (e.g., roads), digital services on the Internet servers (e.g., gambling site), or anything in between". It is argued that face-to-face services are not a part of user experience because there is no user interface involved between the humans. However, product-related services like online trouble-shooting tools for a product, as a service, affect overall UX of the product and also has a separate user experience [1].

Having established the scope of UX, its relationship with usability, and how it differs from other types of experience, a few UX definitions are provided below:

- Alben defined UX as “all the aspects of how people use an interactive product: the way it feels in their hands, how well they understand how it works, how they feel about it while they are using it, how well it serves their purposes, and how well it fits into the entire context in which they are using it” [36].
- Hassenzahl and Tractinsky defined UX as “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.)” [37].
- Hekkert and Schifferstein defined UX as “the awareness of the psychological effects elicited by the interaction with the product” [38].
- Hassenzahl defined UX as “a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service” [39].
- Schulze and Krömker defined UX as “the degree of positive or negative emotions that can be experienced by a particular user in a specific context during and after product use and that motivates for further usage.” [40].
- Nielsen and Norman defined UX as “all aspects of the end-user's interaction with the company, its services, and its products.” [41].
- ISO 9241-210 defined UX as “a person’s perceptions and responses that result from the use and/or anticipated use of a product, system or service” [4].
- Robert and Lesage defined UX as “a multidimensional construct that defines the overall effect over time on the user of interacting with a system and service in a specific context.” [35].

Four factors can be identified when looking at UX definitions: user, product, context (situation) and time. User and product are present in all the definitions; however, some expanded the term “product” by including service, system, or object in their definitions [1, 4, 35, 39]. Researchers and practitioners argue that “customer” should not be used instead of “user” in UX definitions on the ground that UX is not limited to the marketing domain [14]. The importance of the context of use

is highlighted in some definitions like [35, 40] or that of Arhippaninen and Tähti, who defined UX as "people's context-specific experience when interacting with products" [42]. The concept of temporality has also been studied in the literature to highlight the effects of time on UX [33, 43]. As mentioned in the introduction, while ISO 9241-210 includes the anticipated use of a product in its definition, the majority of UX definitions focus on the interaction between the user and the product (i.e., actual use). For instance, Park et al. (2013) did not include brand experience obtained without using the product when evaluating the UX of mobile devices and services [32]. In this research, we adopted the UX definition of Robert and Lesage [35] and focused on the interactive products used by active users.

2.3 UX characteristics

UX characteristics can be summarized as follows according to different definitions and models [35, 37, 44, 45]:

- UX is subjective and individual: It is influenced by what users bring to the interaction (e.g., emotions, expectations, attitudes, interests, moods, knowledge, etc.), as well as their emotional reactions during and after the interaction. As these factors differ from person to person, UX is a unique concept. While individual experiences are unique, the concept of UX can be applied to teams as well.
- UX is context-dependent: UX varies as a result of the specific characteristics of the context, such as the time, place, people, technology, opportunities, incidents.
- UX is dynamic and cumulative: UX evolves with time. Each and every interaction, advertisement, word of mouth, or expectation can enhance or deteriorate UX. Moreover, the importance of UX dimensions may change over time as users gain mastery or develop new needs.
- UX is multidimensional: There are several dimensions related to UX that users encounter when interacting with a product, including physical, psychological, functional, and social dimensions. UX dimensions and sub-dimensions are covered extensively in the succeeding sections.

- UX has various granularity levels: A higher level of granularity of UX involves the user's interaction with an entire system's infrastructure and organization, while a lower level covers a simple component of it.

2.4 UX models

A UX model is a consistent representation of different elements of UX and their relationships. The multidimensional nature of UX has led to the development of various UX models. Depending on the model, specific terms have been used to describe the constituent parts of UX, such as attributes [46], levels [47], components [48], aspects [26], elements [49], qualities [50] and dimensions [35, 45]. In the following, we present some of the UX models that are available in the literature.

2.4.1 UX model of Hassenzahl

Hassenzahl is a well-known researcher in the field of HCI whose UX model provides key elements from the users' and designers' perspective (Figure 2.1). According to this model, designers' aim is to convey an intended product character to users by incorporating a variety of features into their products. Perception of these features varies among users due to their different standards and expectations, resulting in a unique apparent product character for each user. These characters, which may differ from the intended product characters of the designer, include pragmatic and hedonic attributes [46].

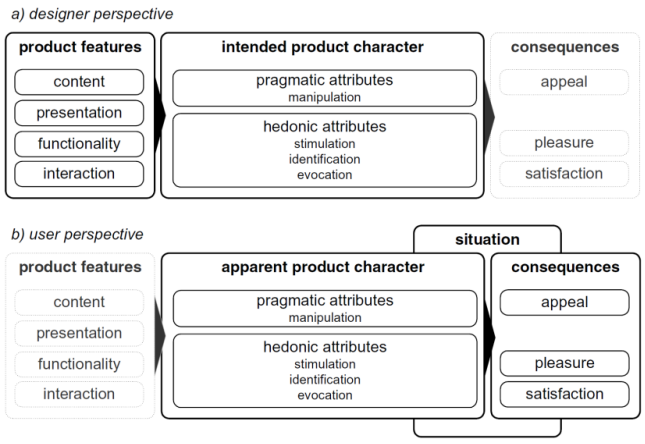

Figure 2.1 UX model of Hassenzahl [46] (reproduced with permission)

Pragmatic attributes deal with achieving behavioral goals that are either externally given or internally generated. There are other names in the literature, such as instrumental, utilitarian, functional, and traditional quality aspects, which denote the same concept [51]. Hedonic attributes pursue accomplishments of a user's psychological goals. Other terms used in the literature include non-instrumental, non-task-related, and affective quality aspects [51]. According to Hassenzahl, hedonic attributes can be divided into three categories: stimulation, identification, and evocation. Hedonic functions of a product can be stimulating by providing users with ways to promote their personal development; helping them identify themselves through usage or possession of a product; and provoking memories attributed to products. Finally, the consequences of users' interactions depend on the usage context and can be reflected in the form of evaluative judgment (appealingness) and emotional reactions (satisfaction and pleasure) [46]. Different products can have different attributes, and a single product can evoke different attributes for the same person in various contexts or over time. Hassenzahl used this UX model to develop his UX evaluation tool called AttrakDiff [6].

2.4.2 UX model of Thüring and Mahlke

Thüring and Mahlke proposed "components of user experience model" (CUE-model) (Figure 2.2). The three main components of UX in the model are instrumental qualities, non-instrumental qualities, and emotional reactions. The first two components are similar to Hassenzahl's pragmatic and hedonic attributes, while the emotional reaction is a result of the perception of the instrumental and non-instrumental qualities of a product. It can be expressed as subjective feelings, physiological reactions, or behaviors. All three components lead to an appraisal of the system and a user decision of future use [48]. Modular evaluation of key components of user experience (meCUE) is the generally applicable UX evaluation tool that was developed later based on the CUE model [8].

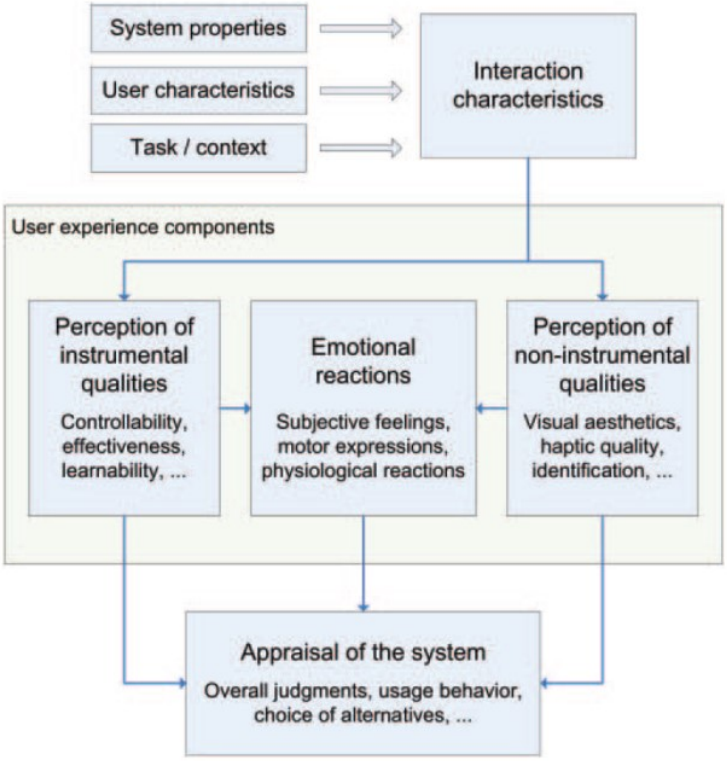

Figure 2.2 UX model of Thüring and Mahlke [48] (reproduced with permission)

2.4.3 UX model of Kort et al.

In their model (Figure 2.3), Kort and colleagues drew inspiration from the works of Wright and McCarthy [52] and Desmet and Hekkert [47]. This model covers both holistic and de-compositional approaches to UX. The holistic approach involves the whole experience before, during, and after interacting with a product. It is presented in the model through the sense making processes of anticipation, connecting, interpreting, reflecting, appropriating, and recounting. The de-compositional approach deals with the components of UX. This model identifies compositional aspects, aesthetics aspects, and aspects of attributing meaning as the three components of UX. Different design elements are applied by designers to convey the intended product aspects to users. Compositional aspect is similar to pragmatic attribute of Hassenzahl's UX model, addressing usability and behavioral characteristics of a product, aesthetic aspect is associated with the look and feel of products, and aspect of meaning is concerned with the higher order goals such as self-realization and personal development of a user [53].

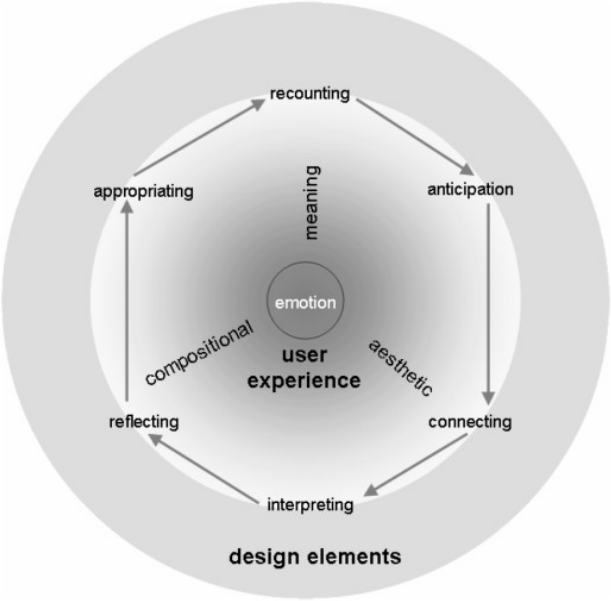

Figure 2.3 UX model of Kort et al. [53] (reproduced with permission)

2.4.4 UX model of Robert and Lesage

Robert and Lesage proposed a model of inputs and outputs of UX (Figure 2.4) that can be seen as the combination of other models. It considers user, product, context, and activity as the main design elements of UX, with aesthetics and sense-making as meta-level elements impacting different aspects of an experience. They have taken into account the role of context in the user-product interaction plus other people, marketing, sale service, after-sale service, and design effects [35, 45].

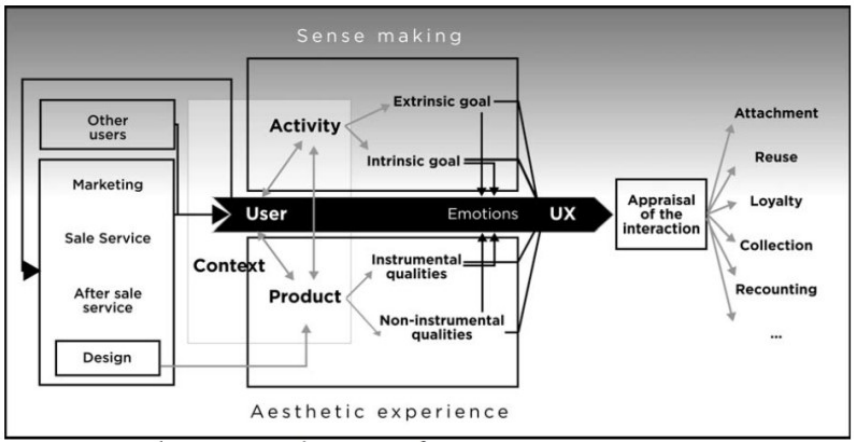

Figure 2.4 Inputs and outputs of UX [35] (reproduced with permission)

These UX models, some of which are the basis of well-known UX evaluation tools, showed the perception of different researchers of the components of UX and their relationship.

2.5 UX dimensions

Defining and categorizing UX dimensions can facilitate the understanding, studying, and evaluating of UX. In our research, we use the terms "dimensions" and "sub-dimensions" as the constituent elements of UX. We consider UX dimensions as the factors that influence the user-product interaction [54] and affect the users' perceptions of their experience with a product, and "sub-dimensions" as attributes [32] or indicators [55] of each dimension. Analyzing the UX tools and models shows different categorizations of UX dimensions. Categorizing UX dimensions into pragmatic and hedonic categories is well documented in the literature, as evidenced by its use in tools such as AttrakDiff, UEQ, and meCUE. According to Table 2.1, the number of sub-dimensions within each category varies among studies.

Table 2.1 UX sub-dimensions covering the pragmatic and hedonic categorization, by studies

Category	**Roto & Rautava [49]**	**AttrakDiff**	**UEQ**	**meCUE**	**Park et al. [32]**	**Orehovački et al. [56]**
Pragmatic	Utility, Usability	Pragmatic	Perspicuity, efficiency, dependability	Usability, usefulness	Usability (simplicity, directness, efficiency, informativeness, flexibility, learnability, user support)	Accessibility, advantageousness, availability, context coverage, compatibility, customizability, dependability, ease of use, effectiveness, efficiency, familiarity, manageability, minimal workload, perspicuity, privacy, scalability, trust
Hedonic	Social value, enjoyment	Identification, stimulation	Novelty, stimulation	Visual aesthetics, status, commitment	Affect, user value (self-satisfaction, pleasure, sociability, customer need, attachment)	Attractiveness, connectivity, novelty, satisfaction, stimulativeness

As a result of observations of, and interviews with people interacting with different products, Robert and Lesage [45] proposed categorizing UX dimensions around two poles: user and product. Their work was further extended by Larouche [57] and Provost [58], resulting in 12 UX dimensions categorized into product and user poles as presented in Table 2.2.

Table 2.2 Categorization of UX dimensions into product and user poles

Pole	Robert & Lesage	Larouche	Provost
Product	Functional	Functional	Functionality (usefulness, functionality options, possibilities)
	Physical	Informational	Usability (simplicity, rapidity, ease of use, efficiency)
		Physical	Informational (presence, relevance, quality)
		Contextual	Physical characteristics (weight, dimensions, size, adjustments)
		Temporal	External characteristics (customer service brand, eco-system)
			Other qualities (accessibility, security, reliability, availability, robustness)
User	Perceptual	Perceptual	Perceptual (aesthetics, presence & quality of multi- media, sense stimulation)
	Cognitive	Cognitive	Cognitive (understanding, concentration, learning, reflection, attention, memory,
	Psychological	Psychological	stimulation)
	Social	Social	Psychological (pleasure/frustration, motivation, expectations, values, evocation, meaning; positive emotions: negative emotions)
			Social (presence of others, quality of interactions, in/dependence from/ to others, obtaining info about others)
			Physical (physical activity, transport, comfort movement, displacement)
			Other personal impacts (productivity, profitability, return on investment)

UX studies have evaluated different UX dimensions depending on the product, context of use, and user type. Following a similar approach to Robert and Lesage, we categorized UX dimensions into more than just pragmatic and hedonic. The result is 13 UX dimension that are also presented in the first article, namely, utility, usability, hedonic, sensory, emotional, informational, social, engagement, cognitive, physical characteristics, impact on body, judgment, overall impression, and others. In the next section, we define these 13 UX dimensions with their related sub-dimensions and evaluation tools.

Usability and utility are the two pragmatic dimensions of UX [59]. **Usability** deals with ease of use and efficiency of a product [60], while **utility** focuses on its function and the qualities of a product that allow users to accomplish their goals with the product. Among the usability sub-dimensions, ease of use, learnability, perspicuity, memorability, and intuitive usage are concerned with how easy a product is to use and reuse. Familiarity, predictability, and consistency are sub-dimensions concerned with the navigation through a product. The other usability sub-dimensions are efficiency and customization, which speed up user interaction; and error tolerance and feedback, which prevent or resolve possible problems. The sub-dimensions of utility contribute to the functionality of a product, including effectiveness, reliability, availability, stability, safety, durability, scalability, and interoperability.

Hedonic dimension is the non-task-related aspect of interaction with a product that differentiates UX from the goal-oriented concept of usability. In our categorization, it includes sub-dimensions

such as novelty, fun, pleasure, enjoyment, stimulation, and evocation, all intertwined with the desirable aspects of using a product. Beauty is among the sub-dimensions that we have separated from the hedonic dimension to cover it under **sensory** dimension. This dimension encompasses the ways in which products are perceived through different senses. A product's visual aesthetics, its sounds and noises, and the tactile experience of the user during an interaction are included in the sensory dimension [3]. There are evaluation tools developed specifically for this dimension, such as VisAWI (visual aesthetics of websites inventory) [61] and those that evaluate it along with other UX dimensions like meCUE.

Depending on the UX model, **emotion** is considered as a UX dimension (e.g., CUE model) or a consequence of the interaction (e.g., UX model of Hassenzahl). This dimension includes the positive and negative emotions that users experience during or after an interaction with a product. Its evaluation can be achieved through tools developed specifically for this purpose, such as PANAS [62] or SAM [63], as well as through UX evaluation tools such as game experience questionnaire (GEQ) [64]and meCUE, that include emotion as a UX dimension.

Recent years have seen an increase in the importance of content as a result of the availability of information provided by both publisher-generated platforms (e.g., Netflix, Disney+, news websites) and user-generated platforms (e.g., YouTube, Instagram, TikTok). The quality, quantity, and trustworthiness of the information presented to users in different formats are evaluated under the **informational** dimension. UEQ+ is among evaluation tools that provide scales to evaluate this dimension.

Using simple communication services like calling or texting, sharing photos on social networks, or challenging friends on a fitness app, social features are growing in different products and software. As a result, we have defined the **social** dimension to include sub-dimensions such as sociability, social acceptance, social relatedness, and conformity. Some evaluation tools like the social-presence module of GEQ [65] go into more detail to evaluate empathy with others, evoked feelings, and behavioral involvement caused by social interactions. Another aspect of the social dimension is the collaboration and cooperation fostered by social interactions (e.g., [66, 67]). Note that identification and similar sub-dimensions like status that are concerned with how a user relates to a product and expresses oneself through its possession or use [46] are not categorized under social dimension in the first article; however, in the third article we have included them to develop social

scales. We have also excluded social behavior exhibited by products like robots or voice-based interactive products (e.g., [68]) from the social dimension.

In addition to gaming consoles and personal computers, smartphones with high performance and efficient processors are attracting more users to play games. Ideally, a game should entertain its players by immersing them in a virtual environment, keeping their attention through a compelling story and exciting gameplay elements, and testing their skills through reasonable challenges. We cover these aspects under the **engagement** dimension. Its sub-dimensions are immersion, presence, flow, absorption, attention, concentration, and involvement.

Learning, creating, and developing knowledge and skills are aspects of interaction with a product covered under the **cognitive** dimension [3] including sub-dimensions such as learning performance, learning helpfulness, and cognition effectiveness. Serious games have been introduced to be used for education, training, and skill development purposes [69], resulting in the development of new UX evaluation tools like the model for the evaluation of educational games (MEEGA+) [70].

Physical characteristics of a product is another important UX dimension that includes aspects such as weight, shape, and dimensions. Wearability, portability, physical suitability, invasiveness, and manageability are its related sub-dimensions. The use of physical and software products can have some impacts on users. Visual discomfort, disorientation, and problem with returning to reality are examples of the effects on the body when using augmented and virtual reality (AR and VR) headsets. We categorized sub-dimensions such as fatigue, workload, tiredness, mental and physical stress, discomfort, and disorientation under the UX dimension called **impact on the body**. There are tools specifically developed to measure this dimension like NASA-TLX (task load index) [71] or FMS (fast motion sickness scale) [72], and other tools like VRNQ (virtual reality neuroscience questionnaire) [73] that incorporate scales to measure the impact on the body for VR applications alongside other UX dimensions

We grouped the sub-dimensions related to the consequences of interaction with a product in the **judgement** dimension. It includes the intention to use, the willingness to recommend a product to others, loyalty, endurability, and trust.

The last dimension of our categorization is the **overall impression**. It addresses the question of whether a product yields a positive or negative experience. This dimension can be evaluated with a single question, like in the meCUE questionnaire or measured with multiple items, like the attractiveness dimension in the UEQ and AttrakDiff questionnaires.

The goal of this categorization was to help identify which UX dimensions are present in UX evaluation tools. We acknowledge that specific application fields or products require custom UX sub-dimensions, such as naturalness for smart homes, social actorship for embodied spoken dialogue systems, and effectance and suspense for interactive digital narratives. We included these field-specific sub-dimensions in the **other** categories.

2.6 UX evaluation

UX evaluation is beneficial to companies by allowing them to determine whether their products meet users' expectations; to diagnose their strengths and weaknesses; and to compare their products to those of their competitors. Evaluation is a key component of almost every design model, to the extent that some consider it the center of the design process [74]. Evaluation and design are closely related, and together they contribute to the improvement of the final product. Similar to the development phase, where users' opinions are valued greatly, evaluation phase can benefit from the users' perspective regarding what UX dimensions are important for a product.

Academic researchers and industry practitioners are not in consensus over UX definition and how it should be evaluated [75]. Most of the UX evaluation methods are developed by academics. It could be due to practitioners' reluctance to publish their methods [76] or their adherence to using usability tools as user experience tools [24]. Moreover, the focus of academic researchers is on theoretical perspectives and qualitative data, whereas UX practitioners prefer practical UX evaluation tools that offer validated measures that are tangible, easy to use and interpret [23]. In the last decade, however, the situation has changed, with an increase in the number of studies using quantitative and mixed methods [17].

UX evaluation is performed at different stages of product development [77]. Researchers have emphasized the necessity of conducting more evaluation studies during the early stages of product development [31, 76]. Evaluations during these stages are conducted on raw ideas, alternative

designs, and certain sections of the entire product [77]. However, some UX dimensions can only be evaluated when users interact with the final product. This is a particular area of interest for businesses to evaluate the long-term UX of products in the market that have been used by different users, as this will allow them to collect as much insight as possible to make necessary improvements to the next version [23]. The following sections present HCI evaluation history, UX evaluation methods, and UX evaluation tools.

2.6.1 History of evaluation in HCI

There have been many stages of evaluation in HCI, beginning with product reliability and then moving on to product performance, user performance, usability, and finally, the user experience. Over the course of time, products, users, and evaluators have evolved [74].

Product reliability was the central focus of the evaluations in the 1940's when the first generation of computers were introduced. The complexity of those machines required professional users to run them and trained evaluators to minimize their failures. Further development of technology shifted the evaluations from reliability to performance in the 1950's. Users and evaluators have also changed from engineers to programmers and computer scientists, who were looking for ways to increase performance by reducing factors such as processing time. In the 1960's, time-sharing processing machines gained popularity and changed users from programmers to people who do non-programming tasks. As a result, user performance evaluation became more interesting for the evaluators [74].

Technology improved drastically during the period of 1970 to 2000, resulting in more novice individuals using different software products. Consequently, more emphasis was put on usability and aspects like learnability and ease of use, in addition to speed and efficiency. Finally, starting in the new millennium, the internet, handheld devices, and technology usage beyond office work made a new phase of evaluation a necessity. UX has emerged to include all aspects of human-computer interaction, specifically hedonic dimensions that have long been neglected [74].

2.6.2 UX evaluation methods

There are three categories of UX evaluation methods: behavioral, subjective, and physiological. Behavioral methods provide objective data on the users' actions and reactions towards a product,

including strategies, performance, and errors. They are employed to observe what users actually do rather than what they claim to do. These data can be quantitative (e.g., task completion time, success rate, error rate) or qualitative (e.g., facial, verbal, postural expressions) and can be collected in real time and remotely. This category of methods is considered to be of medium cost, as it requires software and infrastructure for data collection and analysis. However, behavioral methods cannot explain the reasons behind a user's actions.

Subjective methods provide subjective data on the users' perceptions and appreciation of a product, as well as their workload when using it. Aspects like the user's emotions, expectations, perceptions, judgments, satisfaction, and values can be documented through these methods. In other words, they are about the user's overall experience with a product, so they are usually collected at the end of the experience. Data collection can be done in-person or remotely. This category of methods is considered to be low-cost, and it enables a large number of people to be reached at once because it mostly uses questionnaires. Language, memory constraints, and the biasing effect of social desirability, on the other hand, may be considered shortcomings of subjective measurements [78].

Physiological methods provide objective data by collecting voluntary and involuntary bodily responses during interaction with a product in real-time that overcome language and memory limitations [79]. They collect data using the measurement of electrical activity of the brain, heart rate, pupil dilation, sweating, respiration rate, blood pressure, and muscle tension. They are considered to be expensive because they require investing in tracking equipment and devices, and the evaluator needs data analysis abilities to process the data and make sense of the findings [80, 81]. The devices and sensors used in these methods can be intrusive and uncomfortable for the users who wear them and can make remote evaluation difficult. Lastly, momentary evaluations, typical of psychophysiological evaluations, in most cases do not reliably reflect the real-life experience with a product [82].

2.6.3 UX evaluation tools

Despite the differences, the terms "methods" and "tools" have been used interchangeably in the literature. To clarify, in this research, the **UX method** implies the highest level of evaluation, including behavioral, subjective, and physiological methods that are measured through different data collection methods such as questionnaire, interview, and physiological measure. Whereas **UX**

evaluation tools are instruments that capture users' feedback. For instance, AttrakDiff is a UX evaluation tool in the form of a questionnaire that collects subjective aspects of user interaction with a product. UX evaluation tools and methods have been collected, consolidated, and categorized by different researchers (e.g., [12, 76, 83, 84]). The allaboutux website listed 86 UX evaluation tools/methods based on the study type (e.g., lab, field, online), development phase (e.g., concept, early prototype), type of collected data (e.g., qualitative, quantitative), period of experience (e.g., before usage, long-term), information source (e.g., one user at a time, UX experts), and requirements (e.g., trained researcher, special device). Other categories are recommended such as cost, intrusiveness of the tool, and the UX dimensions that they cover [85].

Considering that the focus of this research is on subjective UX evaluation tools, some of the more established tools are presented below. These tools can be classified as single-dimensional or multidimensional, and as general or specific, depending on the UX dimensions they cover and the context for which they were developed.

2.6.3.1 AttrakDiff

AttrakDiff is a UX evaluation tool that measures multiple UX dimensions and has been used for the evaluation of different products such as websites, smartphones, video games, 3D environments, and learning systems [6]. It consists of four demographic questions and 28 items on a seven-point semantic differential scale measuring both the pragmatic and hedonic dimensions of a product. This subjective tool includes four dimensions: pragmatic (PQ), stimulation (HQ-S), identification (HQ-I), and attractiveness (ATT). The pragmatic dimension describes the usability of a product and the fulfillment of task-oriented goals. The hedonic dimension of stimulation evaluates the extent to which a product fulfills a user's needs to develop one's knowledge and skills, and the hedonic dimension of identification indicates how much a user can identify with a product. Lastly, attractiveness describes the overall value a user gives to a product [6].

AttrakDiff is available online[2] in English and German and is free for UX evaluations with a maximum sample size of 20. It presents the results as in Figure 2.5 in three sections. In section (a),

[2] http://www.attrakdiff.de/

products are categorized based on their ratings on the pragmatic and hedonic axis. This categorization provides a visual representation of differences between two products or two versions of a product. Section (b) provides an average value for each UX dimension, and section (c) gives a detailed description of word-pairs and their ratings.

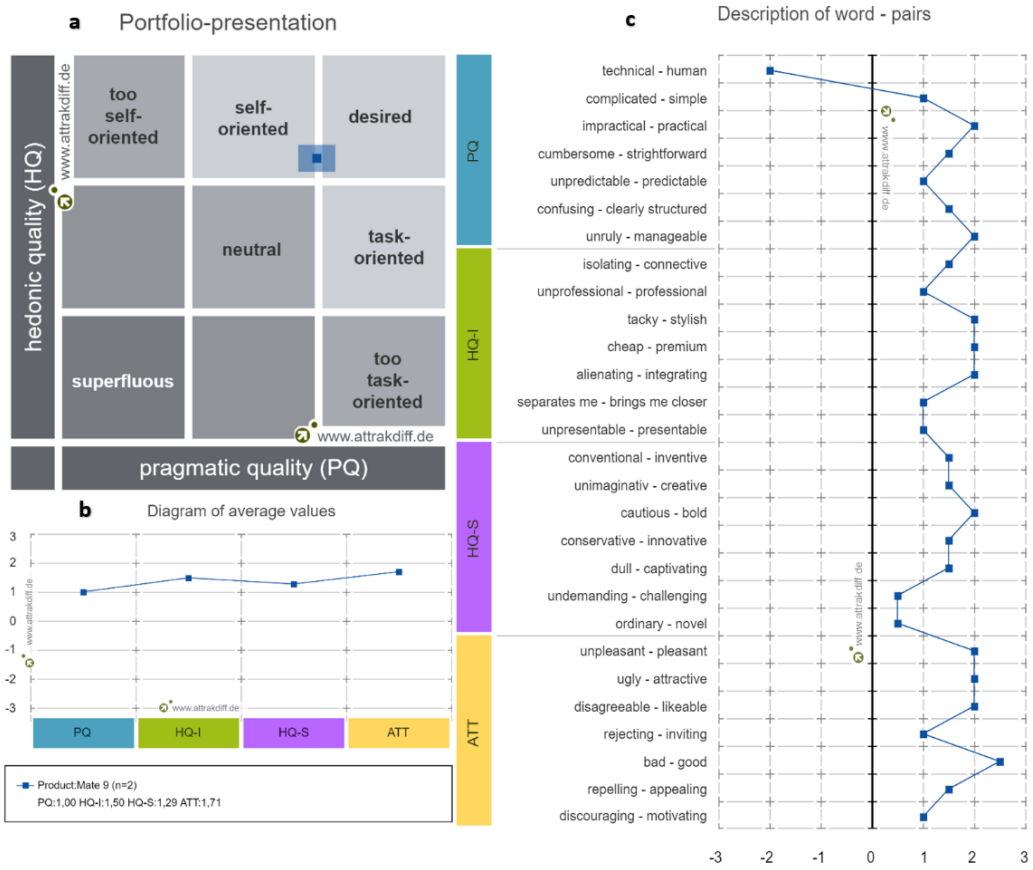

Figure 2.5 Results of the AttrakDiff tool (www.attrakdiff.de)

An abridged version of AttrakDiff (i.e., AttrakDiff2) containing 10 pairs of items was later developed [86]. In this version four pairs represent pragmatic dimension (confusing-structured, impractical-practical, unpredictable-predictable, and complicated-simple), four pairs represent hedonic dimension (dull-captivating, tacky-stylish, cheap-premium, and unimaginative-creative), and two pairs measure attractiveness (bad-good, ugly-beautiful). The abridged version was found to be suitable for differentiating between pragmatic and hedonic dimensions, while being faster to complete [86].

AttrakDiff has the advantage of covering hedonic and pragmatic dimensions, being low cost, usable for a wide range of products, being available online, and providing comparable quantitative data. However, it has been criticized for putting more emphasis on the hedonic dimensions [7] and not being able to rate the importance (weight) of each dimension [87].

2.6.3.2 UEQ

The user experience questionnaire (UEQ) is a multidimensional evaluation tool originally developed for software products that has since been used for different products such as mobile phones, websites, VR headsets, learning systems, and robots [7]. This subjective tool includes six UX dimensions: perspicuity, efficiency, dependability, stimulation, novelty, and attractiveness. The first three dimensions measure pragmatic aspects of UX, the stimulation and novelty measure hedonic aspects, and the attractiveness dimension is similar to that of AttrakDiff, giving an overall judgment of a product. It has 26 items on a seven-point semantic differential scale with four items per dimension except for attractiveness, which has six items (Table 2.3) [7]. UEQ is available online[3] in 36 languages as a spreadsheet file. This tool provides a benchmark that can be used to compare the ratings of a new product with other products that used UEQ (Figure 2.6).

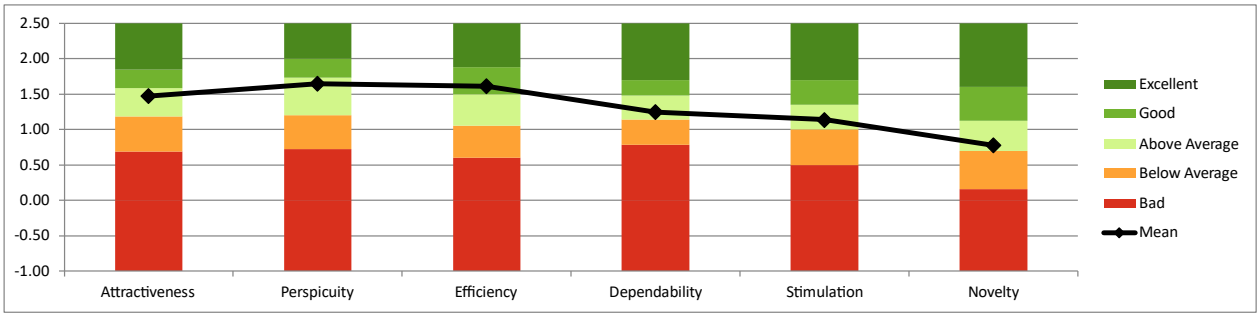

Figure 2.6 UEQ benchmark

A short version of UEQ (UEQ-S) has also been developed, containing eight items (bold in Table 2.3) covering pragmatic and hedonic aspects [88]. Later developments of UEQ introduced a UX

[3] https://www.ueq-online.org/

KPI (Key Performance Indicator) by adding six items measuring the importance of each UX dimension [89].

Table 2.3 UX dimensions and corresponding items of the UEQ evaluation tool

UX dimensions	**Items**	
Attractiveness	attractive – unattractive	pleasing – unlikable
	good – bad	friendly – unfriendly
	pleasant – unpleasant	enjoyable - annoying
Perspicuity	understandable - not understandable	**easy – complicated**
	easy to learn – difficult to learn	**clear - confusing**
Efficiency	fast – slow	organized – cluttered
	efficient – inefficient	practical - impractical
Dependability	predictable – unpredictable	secure – not secure
	supportive – obstructive	meets expectation – does not meet expectation
Stimulation	**interesting – not-interesting**	valuable – inferior
	exciting – boring	motivating - demotivating
Novelty	creative – dull	innovative – conservative
	inventive – conventional	**leading edge - usual**

Main advantages of UEQ are that it is quick, easy to use, available online in different languages, it provides comparable quantitative data, and considers the importance of each UX dimension. However, the benchmark feature of this tool is not robust in that it does not differentiate between various products that have been evaluated by UEQ, although future studies are expected to separate benchmarks by product category [90].

2.6.3.3 meCUE

Modular evaluation of key Components of User Experience (meCUE) is another generally applicable multidimensional evaluation tool. It has been developed based on the CUE model [48] and was first introduced in German [91] and then translated into English [92]. It consists of 34 statements evaluated on a seven-point Likert scale. It has a modular structure comprising 1) product perceptions, 2) emotions, 3) consequences of use, and 4) overall UX modules. The evaluators select the modules that best apply to the UX dimensions they are interested in. MeCUE is a validated tool that is compared with other well-known tools like AttrakDiff, UEQ, SAM, and PANAS [8]. MeCUE was criticized for being unsuitable for business-oriented applications where hedonic

dimensions are viewed as irrelevant [93]. As a result, the second version of the tool (meCUE 2.0) splits the first module of product perceptions into two separate modules: 1) perception of instrumental product qualities and 2) perception of non-instrumental qualities [94]. Instrumental qualities include usability and usefulness, while non-instrumental qualities include aesthetics, status, and commitment [94].

meCUE is available online[4] in English and German as a spreadsheet file. Compared to UEQ and AttrakDiff, it has the advantage of covering emotions, so it does not require the use of another tool to evaluate this dimension. It is leaner than the other two questionnaires because of its modularity. However, the responsibility of selecting the correct modules is on the evaluators, who need to make sure that items appropriately fit with the context. Unlike UEQ, meCUE does not provide importance ratings for each module.

2.6.3.4 UEQ+

User Experience Questionnaire plus (UEQ+) is a modular framework based on UEQ, including 20 UX dimensions that can be used for the evaluation of different products [50]. These dimensions are attractiveness, perspicuity, efficiency, dependability, stimulation, novelty, aesthetics, adaptability, usefulness, intuitive usage, value, trustworthiness of content, quality of content, trust, clarity, haptics, acoustics, response behavior, response quality, and comprehensibility. Each dimension consists of four items measuring the dimension and a single item that determines the importance of that dimension for the UX evaluation of the product. The evaluation of an item is done on a 7-point Likert scale with two semantic differential anchor points.

UEQ+ is available online[5] in more than 20 languages as a spreadsheet file. The modularity of the UEQ+ enables it to be used in different contexts, allowing the evaluator the flexibility to tailor the tool and evaluate the set of UX dimensions most relevant to the study. The consistency of UEQ+ simplifies the UX evaluation compared to using a combination of UX evaluation tools, each with different rating scales and possible overlapping dimensions. Other advantages of this tool include

---

[4] www.mecue.de

[5] www.ueqplus.ueq-research.org

being comprehensive, flexible, easy to use, low-cost, and rapid. Moreover, researchers can construct new UX dimensions and add them to the UEQ+ framework. For instance, Boos and Brau [95] added two UX dimensions (haptic and acoustic) for household devices, and Klein et al. [96] developed three new dimensions (response behavior, response quality, and comprehensibility) for the UX evaluation of voice assistants.

2.6.3.5 GEQ

Game Experience Questionnaire (GEQ) is a case-specific, multidimensional, and modular evaluation tool. It consists of three modules: 1) core, 2) social presence, and 3) post-game (see Table 2.4). These modules are provided to players after their gaming session to evaluate different aspects of their interaction [64]. In addition, a shorter version, known as in-game GEQ (iGEQ), has been developed to be used more than once during a gaming session.

Table 2.4 GEQ modules and corresponding UX dimensions

Modules	**Sub-dimensions**	
Core	Competence	Challenge
	Immersion	Negative affect
	Flow	Positive affect
	Tension/Annoyance	
Social presence	Empathy	Behavioral involvement
	Negative feelings	
Post-game	Positive experience	Tiredness
	Negative experience	Returning to reality

The advantage of this tool is that it includes individual and group gaming sessions, it is applicable in laboratory and field studies, and provides the option to choose different modules. The difficulty of answering some items after a short playing time in the laboratory setting is pointed out as its main disadvantage [83].

2.6.3.6 PANAS

The Positive and Negative Affect Scale (PANAS) is a single-dimensional self-reported questionnaire that uses two 10-item scales to assess users' positive and negative emotions [62]. It has been used in conjunction with other UX evaluation tools to capture users’ emotions when using a product. Other versions of this tool include PANAS for children (PANAS-C), a short form

(PANAS-SF), an international short form (I-PANAS-SF), and an expanded form (PANAS-X). Being a validated tool that has been used in different domains is its main advantage. However, the absence of any explanation for items can be counted as a disadvantage, as it may cause different interpretations on the part of users and influence the results.

2.7 Synthesis

With the increasing number of products, services, and systems we interact with in our daily lives, the importance of UX has grown significantly. There is a long list of UX evaluation tools available in the literature that can be used in different situations. UX subjective evaluation tools are the most commonly used tools due to their low-cost and ease-of-use. We saw that the UX dimensions evaluated by each of these tools can vary greatly. Despite covering different pragmatic and hedonic aspects by these tools, we found a lack of attention to evaluate the social aspects of the user experience. This is surprising, as online social network platforms are commonplace and social features are being added to existing products to differentiate them from the competition (e.g., being able to share a cart with friends on an online shopping website). Knowing how socially attractive a product is, and how its social features influence the overall UX are of growing importance and deserve more attention from the research community.

We have also found that more recent UX subjective evaluation tools are built modularly. This is the case of UEQ+, which provides the option to develop and add of new UX dimension scales without modifying the existing, validated scales. If UEQ+ follows the same growing adoption of UEQ from UX researchers and practitioners, it will be a promising avenue to investigate different UX dimensions, including the social dimensions.

CHAPTER 3 THESIS ORGANIZATION

This thesis is organized in eight chapters. Chapter 1 provided the context, problems, and objectives of this research. Chapter 2 presented the review of the UX literature. Chapter 3 provides the organization of the three articles in relation to the research objectives.

Chapter 4 (article 1) presents a SLR on UX dimensions and UX subjective evaluation tools during the time period from 2010 to 2021. It provides the state of the art in UX regarding the countries with the most contributions to this field, products evaluated, evaluation time, duration of product use before evaluation, evaluation methods, evaluation tools, and the UX dimensions that were covered. It fulfills the first objective by revealing the new trend of modularity in general UX evaluation tools (e.g., meCUE, UEQ+) as well as the underrepresentation of informational, social, physical, and cognitive dimensions in the current UX evaluation tools.

Building on the results of the SLR, chapter 5 (article 2) investigates the importance of 21 UX dimensions for 15 different product categories. It accomplishes the second objective by providing a list of important UX dimensions for different product categories that can facilitate the use of modular tools. It compares the importance ratings given by end users and UX experts. It also provides a cross-cultural comparison of the results obtained from Canada (this research) with those done by other researchers in Germany and Indonesia.

Chapter 6 (article 3) delves deeper into the social dimension of UX and accomplish the third objective of this thesis by developing and validating social scales that could be added to the UEQ+ modular framework. It starts with defining the initial pool of items. Next, UX experts reviewed them in an online workshop and finalized the candidate items for factor analysis. The first questionnaire collected the UX of products having social features, the results of which underwent exploratory factor analysis resulting in the development of four social scales. The final step included collecting the UX of three products with varying levels of social aspects to validate the social scales.

Chapter 7 presents a general discussion of the three articles combined. Finally, chapter 8 provides concluding remarks, addresses the fulfillment of the thesis' objectives, and recommends possible directions for future research.

CHAPTER 4 ARTICLE 1: SYSTEMATIC LITERATURE REVIEW ON UX DIMENSIONS AND UX SUBJECTIVE EVALUATION TOOLS

Ehsan Mortazavi, Philippe Doyon-Poulin, Daniel Imbeau, Mitra Taraghi, Jean-Marc Robert.

This article was submitted to Interacting with Computers Journal on March 23, 2022.

Abstract

The quality of the User Experience (UX) with systems, products, and services is now considered an indispensable part of success in the market. Users' expectations have increased in such a way that mere usability is no longer sufficient in a large range of situations. In this study, we conducted a systematic literature review on UX subjective evaluation tools and the UX dimensions covering the period of 2010-2021 with an initial sample of 3831 publications, 325 of which were selected for the final analysis, in order to provide researchers and practitioners with the recent changes in the field of UX. Main results showed that 104 different tools are available for UX evaluation, they can be classified as general or domain-specific, applicable for a wide variety of products, and in total covering more than 300 UX dimensions. Our categorization of UX dimensions under 13 main dimensions (e.g., usability, utility, hedonic, emotion, sensory, etc.) showed that the informational, social, physical, and cognitive dimensions appeared to be less frequently present in current tools. We argue that these four dimensions deserve more space in UX tools. Having a high number of UX evaluation tools can be confusing for evaluators, and they need some guidance for selecting and combining tools. We discovered that modularity is the new trend in the development of UX evaluation tools (e.g., meCUE, UEQ+) with the benefits of being comprehensive, flexible, easy to use, low-cost, and rapid, avoiding overlapping of dimensions and providing comparability through the use of a similar format and rating scale. Finally, the need for having a comprehensive evaluation tool requires updating the set of included dimensions to accommodate for new generations of products and technologies.

Keywords: User experience, UX evaluation tool, UX dimension, systematic literature review, SLR

4.1 Introduction

The development of technology has enabled organizations to introduce systems, products, or services[6] that provide new features, engage more senses, and create new experiences for users. As technology evolved from the first generation of computers in the 1940s to today's ubiquitous devices, users shifted from engineers and highly trained professionals to general people [74]. The field of user experience (UX) has also grown and evolved since its introduction more than two decades ago to address the changes in users and industries.

The number of research initiatives in the field of UX has increased during this period. UX definition, UX dimensions, evaluation tools and methods are among the subjects that have been investigated in different studies. ISO 9241-210 defines UX as the "*user's perceptions and responses that result from the use and/or anticipated use of a system, product or service*" [4]. Although criticized [11], this is the most widely used definition of UX. Researchers in UX come from a variety of disciplines, including computer science, ergonomics, psychology, anthropology, graphic arts, software design, market research, and branding [25]. Efforts have been made to reach a universal definition of UX [97], based on the researchers' disciplines and perspectives [1] [98], but no consensus was reached.

UX dimensions are the different factors that influence the user-product interaction [54] and have impact on the users' perceptions of their experience with the product. Hedonic and pragmatic dimensions are the two major groups of UX dimensions [46]. Fulfillment of psychological needs, pleasure, and aesthetics are covered by hedonic dimensions, while functional or utilitarian aspects are addressed by pragmatic dimensions [46]. Emotional, sensory, informational, social, and physical aspects are other UX dimensions mentioned in recent studies [44, 54]. The UX dimensions considered, and their evaluation tools tend to change based on the researchers' disciplines and the types of products they are interested in for the UX. New UX dimensions emerge with the introduction of new products [99] and new technologies; thus, evaluation methods and tools should be flexible to accommodate these changes.

---

[6] For the sake of brevity, only the term product is used in the following text.

UX evaluation can be done either by experts or end users. Experts observe or anticipate users' interaction with products, use heuristics, expert review, and cognitive walkthrough [100], while end users can fill out questionnaires, participate in interviews, or think aloud while interacting with a product as methods to report on their UX. According to ISO 9241, "*UX includes users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur before, during and after the use of a product, system or service*" [4]. It confirms that the three categories of evaluation methods (i.e., behavioral, subjective, and physiological) are all applicable in the UX evaluation.

- Behavioral methods provide objective data on the users' actions and reactions towards a product, including strategies, performance, and errors. They are employed to observe what users actually do rather than what they claim to do. These data can be quantitative (e.g., task completion time, success rate, error rate) or qualitative (e.g., facial, verbal, postural expressions) and can be collected in real time and remotely. This category of method is considered to be of medium cost. However, behavioral methods cannot explain the reasons behind a user's actions.
- Subjective methods provide subjective data on the users' perceptions and appreciation of the product, and their workload when using the product. Aspects like the user's emotions, expectations, perceptions, judgments, satisfaction and values can be documented through these methods. In other words, they are about the user's overall experience with a product, so they are usually collected at the end of the experience, remotely or not. This category of methods is considered to be of low-cost, and it enables for a large number of people to be reached at once because it mostly uses questionnaires. However, language, memory restrictions and the biasing effect of social desirability may be counted as shortcomings of subjective measurements [78].
- Physiological methods provide objective data by collecting voluntary and involuntary bodily responses during interaction with a product in real-time that overcome language and memory limitations [79]. They collect data using the measurement of electrical activity of the brain, heart rate, pupil dilation, sweating, respiration rate, blood pressure, and muscle tension. Nonetheless, they are considered to be of high cost because they

require investing in tracking equipment and devices, and data analysis abilities to process the data and make sense of the findings [80, 81]. The devices and sensors used in these methods can be intrusive and uncomfortable for the users who wear them and make remote evaluation difficult. Lastly, momentary evaluations, typical of psychophysiological evaluations, in most cases do not reliably reflect the real-life experience with a product [82]

The COVID pandemic has imposed new constraints and costs for laboratory and field studies due to health measures aimed at protecting the evaluators and the participants. Overall, these three methods do not take precedence over one another, rather, it is suggested to triangulate objective and subjective data provided by different methods to gain a holistic view of UX [80].

Comprehensive UX studies have taken various forms such as snowballing study, systematic literature review (SLR), and systematic mapping study to explain and classify different aspects of this field, like UX dimensions, UX evaluation tools and methods, evaluated products, types of collected data and more. In the years 2000-2010, researchers focused on the UX definition [1, 98] and the available tools and methods for UX evaluation [76, 101], and positioned the UX as a new concept that goes beyond usability. Roto et al. investigated 30 UX evaluation methods with professionals in academic and industrial contexts [101]. Their work was completed by Vermeeren et al. who collected 96 UX evaluation methods from literature reviews, workshops, special interest groups and online surveys [76]. Their results are published on the allaboutux.com website, listing 86 UX evaluation methods and their characteristics such as source of information, location, product development phase, and type of collected data, to name a few. The systematic literature review of Bargas-Avila and Hornbaek [16] is the main source for the following reviews in the field of UX [5, 84, 102]. They analyzed the empirical UX studies regarding the products, UX dimensions, and methodologies during the time period from 2005 to 2009 [16]. Their findings showed that the context of use shifted from work to leisure, with evaluations in open use situations with consumer products. The most frequently evaluated UX dimensions were emotions, enjoyment and aesthetics, and studies mostly used qualitative methods [16].

From 2010 onward we observed the emergence of UX literature reviews on specific products like ubiquitous computing systems [103], information driven website [104], language management systems [105], interactive digital narratives [106], natural user interface [107], and Virtual Reality

(VR) systems [108]. Literature reviews investigating the state of the art in UX were published as well. Rivero and Conte [109] performed a systematic mapping study to discover the UX technologies (tools, methods, techniques) of software applications during the 2010-2015 time frame, but their work was limited to only one database (Scopus). The systematic review of Maia and Furtado [82] revealed that psychophysiological evaluation methods have not been used extensively in UX studies, the number of studies with long term UX evaluation was low, and subjective UX evaluation tools like questionnaires were used more than all the other tools. Their work, however, was based on 25 papers only. The SLR by Robinson et al. [5, 17] followed the work of Bargas-Avila and Hornbaek by extending the time period from 2000 to 2016. In addition, they also explored empirical studies on UX regarding the use of research questions or hypotheses. They covered 400 empirical studies extracted from one database (google scholar). The work of Zarour and Alharbi [110] reviewed UX publications with a focus on software and technology during the 2005-2015 timeframe and provided a better picture of UX studies including brand, user and technology experiences. Covering the period of 2010 to 2016, Pettersson et al. [102] investigated the triangulation of methods in UX studies published in eight ACM conferences. The most recent literature reviews are the work of Darin et al. [111] and Diaz et al. [9]. Darin and colleagues [111] used a snowballing procedure on the Vermeeren et al. list of UX evaluation methods and provided an updated list of UX instruments developed up to 2018. However, their review missed recent UX subjective evaluation tools such as meCUE[7], MUX[8] and UEQ[9]+. Diaz and colleagues [9] only reviewed three UX questionnaires (i.e., AttrakDiff, UEQ[10] and meCUE).

The state of the art on subjective UX evaluation tools is lacking, since previous SLRs used a limited number of databases [5, 82, 102], the studied time frame ended more than 5 years ago [5, 110], and many studies included focused on a subset of subjective questionnaires [9]. Since subjective questionnaires are the evaluation tool used the most in UX studies [16], there is a clear need to

---

[7] Modular evaluation of key components of user experience

[8] Mobile User eXperience

[9] User Experience Questionnaire plus

10 User experience questionnaire

provide a critical state of the art on the topic to offer up to date information regarding the UX dimensions measured by existing questionnaires.

In this study, we conducted a SLR on UX subjective evaluation tools and the UX dimensions considered. We reviewed literature from 2010 to 2021 using four databases. We expected that conducting a SLR with more databases compared to previous SLRs would yield more reliable and comprehensive results. We focused on the UX dimensions evaluated by these tools and the products they were applied to. This is of interest since recent products introduced new UX dimensions that were not considered in previous research [99], and investigating which products have been studied in recent years can help researchers update their tools and methods.

4.2 Methodology

The present SLR was conducted based on the procedure proposed by Kitchenham and Charters [112], that includes three main phases:

- Planning the review, which includes establishing the research questions and determining the search strategy.
- Conducting the review, which includes performing the search and refining the primary results based on the selection criteria.
- Reporting the review.

4.2.1 Planning the review

In this SLR, we reviewed UX empirical studies with interactive products. By empirical, we mean studies in which a real interaction took place between the user and a product (so the study presents user data). The SLR does not include studies with imagined products, expert reviews (e.g., heuristics), nor observer data.

Research questions: we formulated three research questions:

RQ1- Which products, in which stage of development, where, when and how, were evaluated in the UX empirical studies?

RQ2- Which UX subjective evaluation tools were used?

RQ3- Which UX dimensions were covered?

Database selection: We selected four databases for the SLR. ACM Digital Library is a rich source in the field of Human-Computer Interaction, covering established conferences and journals like CHI[11] and TOCHI[12]. IEEE Xplore Digital Library and Web of Science were also incorporated to cover a wider range of scientific literature. Robinson and Lanius [17] showed that in addition to HCI and computer and information science, engineering journals also published UX studies. Therefore, we added Engineering Village as a comprehensive and multi-disciplinary database to ensure better coverage.

Search terms: A set of specific keywords were selected based on the research questions. The terms "User experience" and its shorter form "UX" were the main ones. A first group of terms that refer to UX dimensions in the literature like component, attribute, aspect, indicator, factor, dimension and element were also included. Additionally, evaluation, measurement and tool were the other group of terms applied in this SLR. The search strings for each database were formulated with the help of a librarian specialist and are provided in Appendix A.

Inclusion criteria:

- The paper answers one of the research questions.
- The publication type is a journal or a conference paper.
- The paper has been published between January 2010 and April 2021.

Exclusion criteria:

- The paper does not evaluate the UX nor address its dimensions.
- The paper is not related to the domain of user experience.
- The evaluation tool/method used is not subjective (e.g., physiological, behavioral).
- The evaluation is done by someone other than the user (e.g., expert, observer).

---

[11] Conference on Human Factors in Computing Systems

[12] Transactions on Computer-Human Interaction

- The product evaluated is not interactive.
- There is no interaction between the user and the product.
- The paper is written in a language other than English.
- The full text is not available.

4.2.2 Conducting the review

The scientific papers were searched in the subject, title and abstract fields for Web of Science and Engineering Village databases, in the abstract field for ACM Digital Library, and in all fields for Web of Science repositories. Search strings varied amongst databases because of different search functionalities, however they all used the same terms. The search was performed in April 2021 and produced 3831 papers (see Figure 4.1). The results were imported to Endnote X9 for further evaluation. First, 1357 duplicates and papers that appeared in more than one database were removed. Second, the remaining 2474 publications were filtered to remove 75 papers that were not in English or had a format other than that of a journal article or conference paper (e.g., poster or abstract). The third step required reading the abstracts of 2399 papers to evaluate them based on the inclusion criteria. A small number of papers (31) in other fields like aeronautics, network, physics, and material science were removed because the term “UX” was used with a different meaning. An additional 868 papers in which UX was not the main subject of the paper and was only mentioned in the abstract or keywords were also removed. The full texts of 50 papers were not available and thus they had to be removed. Overall, 949 papers were removed in this step.

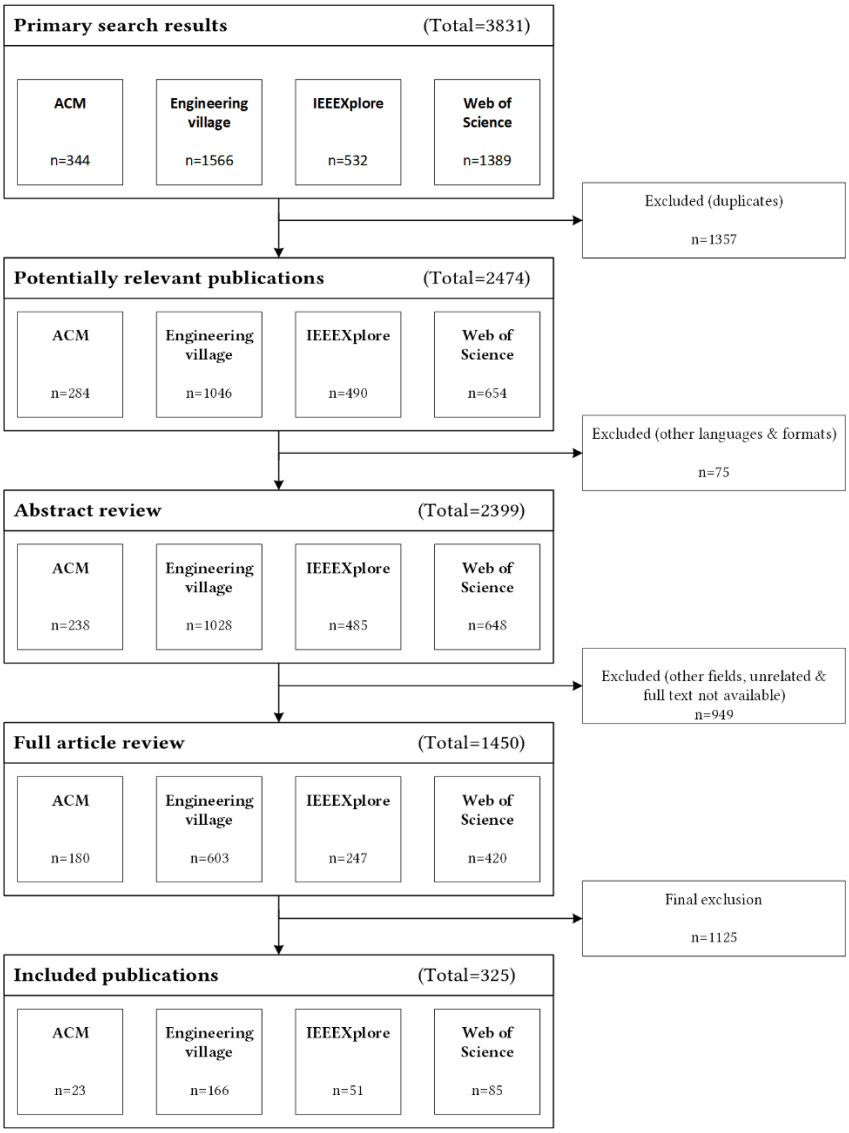

Figure 4.1 SLR flow chart

Fourth, the full texts of the 1450 remaining papers were downloaded for further analysis. By applying the selection criteria, 1125 papers were excluded (i.e., subjective tools not applied; evaluation done by an observer; think aloud, interview, physiological or performance measures used; mentioned UX but evaluated usability or technology acceptance; did not mention the dimensions used for the evaluation of UX). The process yielded a corpus of 325 papers, 310 of which were UX empirical studies, and 15 were just focused on method, model, and tool development without empirical data. The complete list of 325 articles reviewed is available in the supplemental material section on the Journal's website.

4.2.3 Grouping UX dimensions

The first author investigated the UX tools used in the selected studies and extracted their dimensions to determine which UX dimensions were covered, making a list of 324 UX dimensions. Then, two cognitive ergonomics PhD candidates removed the dimensions that had the same concept but different names. Next, a full professor with 40 years of experience in UX research, joined the team for the classification of UX dimensions during two online meetings. The classification was based on the work of Robert [3] in which 8 main dimensions were identified, functional, usability, informational, physical characteristics, sensory, cognitive, psychological, social, and physical (user-related). The odd number of subject matter experts helped to break the tie in case of disagreement. The final classification included 13 main dimensions as presented in the results section.

4.3 Results

This section presents the analysis of the 325 papers selected, pointing out the year of publication, type of product evaluated, field of use, study type, evaluation time, duration of interaction, types of collected data, data collection tools and methods, and the dimensions of UX studied.

4.3.1 Year and country of publication

The number of publications per year addressing UX dimensions and subjective evaluation tools that met our criteria for the SRL has shown a steep increase since 2017 (Figure 4.2). Note that results from 2021 are not representative of the whole year since the search considered papers published until April 2021.

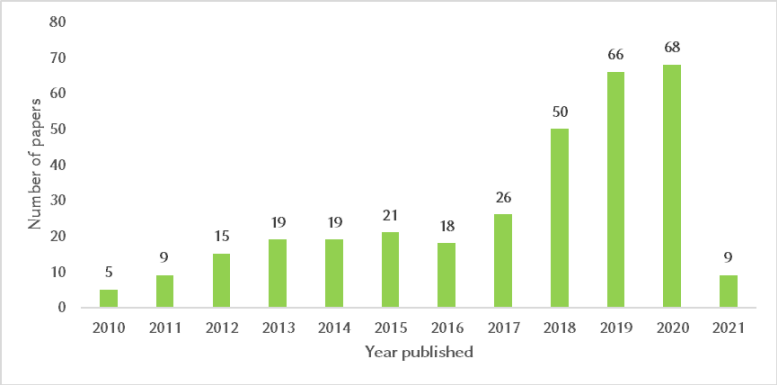

Figure 4.2 Distribution of UX publications per year (n=325)

The distribution of publications per country was determined based on the affiliation of the first author. This SLR covers publications from 47 countries among which those with more than 5 papers are shown in Figure 4.3. Germany, home of pioneering UX evaluation tools like AttrakDiff [6] and UEQ [7], had the greatest number of publications on the subject, followed by two Asian countries. Similarly, among continents Europe had the greatest number of papers with 183, followed by Asia with 89 publications.

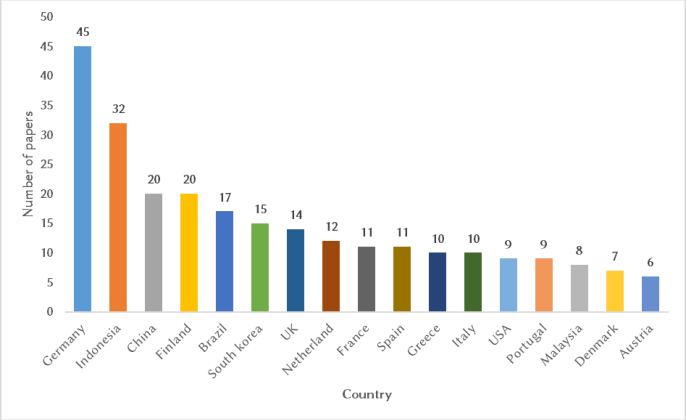

Figure 4.3 Distribution of UX publications per country (n=325)

4.3.2 Products evaluated

The most frequent type of products evaluated with subjective assessment tools (Table 4.1) were software applications such as mobile apps, computer software, and digital games, and accounted for 49.7%. The term “system” in the table corresponds to a combination of software and hardware, such as VR training systems [113], collaborative robots [114], and multimodal tools for

presentation [115]. The group “Other” includes products like virtual assistants [96], spine posture monitor wearable instruments [116], and Brain Computer Interaction (BCI) interfaces [117], to name a few. In some of the studies more than one type of product was evaluated, such that the total number of products evaluated amounted to 332.

Table 4.1 Number of evaluated products in the SLR (n=332)

Product	Frequency	Percentage
Software application	165	49.7
System	67	20.2
Website	44	13.3
Other	56	16.8

Overall, classifying the products into two main groups of software and hardware highlights the dominance of former (Table 4.2).

Table 4.2 Number of product types in the SLR (n=332)

Product type	Frequency	Percentage
Software	236	71.1
Hardware	11	3.3
Both	85	25.6

In the studies, users were required to evaluate their UX with a prototype (low or high fidelity) or with a finished product, depending on the stage of development and the goals of the researchers. UX dimensions such as aesthetics can only be accurately measured after interacting with a final product. Results in Table 4.3 show that the majority of UX studies evaluated final products.

Table 4.3 Number of studies per stage of development of evaluated products (n=332)

Stage of development	Frequency	Percentage
Final product	240	72.3
Prototype	86	25.9
Both	6	1.8

4.3.3 Field of use

Bargas-Avilla and Hornbaek [16] classified products into leisure, work or mixed, but UX has been used in a variety of fields. In this study, we grouped them into entertainment, learning & education, health, work & training, tourism & museum, and others as presented in Table 4.4. The results show

that UX studies are no longer limited to leisure products, but rather have expanded to other fields such as new technologies like AR and VR, which have broadened users' experiences.

Table 4.4 Number of papers per field of use in the SLR (n=310)

Field of use	Frequency	Percentage
Entertainment	59	19.0
Learning & Education	58	18.7
Health	28	9.0
Work & Training	20	6.5
Tourism & museum	11	3.5
Others	134	43.2

4.3.4 Study type

Depending on the product, stage of development, cost, and other considerations, the UX evaluation may be conducted in the field, the laboratory, both, and online. Table 4.5 shows that more than half (51.8 %) of the studies were conducted in a laboratory-setting. Experiments that require specific conditions or use early prototypes are examples of laboratory-based evaluation studies (e.g., head-mounted display integrated with a pulley system to simulate the forces on the user's face in a VR game) [118]. Field studies represented 20.2% of the studies reviewed and included, for instance, the evaluation of a personal breathalyzer for alcohol consumption by college students over a period of two weeks [119] or an AR app in a museum [120].

Table 4.5 Number of studies per study types in the SLR (n=332)

Study type	Frequency	Percentage
Laboratory	172	51.8
Field	67	20.2
Online	29	8.7
Both	14	4.2
Not specified	50	15.1

4.3.5 Evaluation time

UX can be evaluated at different points in time. Anticipated UX is evaluated before using a product, momentary UX is recorded during its use, episodic UX is measured after its use, and cumulative UX considers multiple durations of use [33]. Table 4.6 shows that most studies evaluated UX after the interaction with the product. Questionnaires, rating scales, interviews, and pictorial measures

are subjective evaluation tools used after interaction with a product, whereas physiological methods like EEG[13], HR[14], EDA[15] or eye tracking are used during the interaction. Because the focus of this SLR was on subjective evaluation tools and methods, we expected to find more studies with UX evaluations done after user interaction with the product.

Table 4.6 Number of studies per evaluation time (n=332)

Evaluation time	Frequency	Percentage
After the interaction	271	81.6
During and after the interaction	40	12.0
Before and after the interaction	19	5.7
Before, during, and after the interaction	2	0.6

4.3.6 Duration of interaction

UX is dynamic and it evolves with time [3]. Therefore, it is important to know how long a user has been interacting with a product before evaluating the associated UX. In Table 4.7, the term "short" refers to an interaction lasting less than a couple of hours, "long" refers to duration of use that range from days to months, and "cumulative" refers to products owned by users and that have been used regularly over long periods. Most studies reviewed (69.9%) evaluated UX after less than a couple of hours of interaction.

Table 4.7 Number of studies per duration of interaction (n=332)

Duration of interaction	Frequency	Percentage
Short	232	69.9
Long	36	10.8
Cumulative	27	8.1
Short & long	4	1.2
Not available	33	9.9

[13] Electroencephalography

[14] Heart rate

[15] Electrodermal activity

4.3.7 Types of interaction

In UX evaluation studies, users are either free to interact with a product or are required to complete a set of specific tasks. Free interaction allows users to familiarize themselves with a product and reflects the overall UX, whereas performing a set of specific tasks better captures their experience with different parts of the product. As a result, combining the two will yield better results. Measuring user performance in a specific task set, having limited time for the evaluation, and the stage of development of the product (i.e., final product or prototype) are among factors that researchers and practitioners need to consider when deciding on the type of interaction. Surprisingly, we found that both categories were equally represented in the studies reviewed (Table 4.8).

Table 4.8 Number of studies per type of interaction (n=332)

Type of interaction	Frequency	Percentage
Task set	154	46.4
Free	140	42.2
Both	20	6.0
Not available	18	5.4

4.3.8 Types of data collected

There are two types of data collected in the publications reviewed: qualitative and quantitative. While the former can provide more in-depth information about an experience, the latter can usually be collected in a shorter time. Table 4.9 shows that most studies used only quantitative data (56.1%) or a combination of both quantitative and qualitative data (42.6%).

Table 4.9 Number of papers per type of collected data (n=310)

Type of data collected	Frequency	Percentage
Quantitative	174	56.1
Qualitative	4	1.3
Both	132	42.6

4.3.9 Data collection methods

Table 4.10 shows the methods used to collect subjective data on UX and their frequency. As expected, questionnaires and rating scales have been used more than any other methods. The other

methods used were interview, think aloud, pictorial (e.g., SAM[16], PrEmo), drawing (e.g., iScale, UX curve), writing (e.g., sentence completion, diary), card and board (e.g., MAX[17], reaction cards) and tactile (e.g., SEI[18], TACTUX[19]).

Table 4.10 UX evaluation methods used in the studies covered by this SLR (n=594)

Data collection method	Frequency	Percentage
Questionnaire/Scale	426	71.7
Interview	72	12.1
Think-aloud	38	6.4
Pictorial	30	5.1
Drawing	11	1.9
Writing	9	1.5
Card & Board	6	1.0
Tactile	2	0.3

4.3.10 Evaluation tools in UX studies

In our review, we found 104 different tools used separately or in combination with others to measure UX and allotted them to two categories based on the UX dimensions they measured. You can find an extract of these tools in Table 4.11. The first category includes tools that evaluate multiple dimensions of UX within a single questionnaire (e.g., AttrakDiff, UEQ, and meCUE). The second category includes tools that target a specific dimension of UX like usability (e.g., SUS[20], CSUQ[21]), emotion (e.g., SAM, PANAS[22]), engagement (e.g., UES[23], GEQ[24]), presence (e.g., PQ[25],

[16] Self-Assessment Manikin

[17] Method for the Assessment of eXperience

[18] Sensual Evaluation Instrument

[19] Tactile user experience assessment board

[20] System Usability Scale

[21] Computer System Usability Questionnaire

[22] Positive and Negative Affect Scale

[23] User Engagement Scale

[24] Game Engagement Questionnaire

[25] Presence Questionnaire

SUS[26]), etc. Given the multidimensionality of UX, the use of a combination of these questionnaires appears appropriate.

UEQ has been used more than any other UX evaluation tools, followed by AttrakDiff, another multidimensional UX questionnaire (see Table 4.12). Among dimension-specific tools, usability was measured most often with SUS, emotion with SAM, workload with NASA-TLX and simulator sickness with SSQ.

Table 4.11 List of tools used in UX studies per UX dimension or subject

Dimension/subject	**Tool**
Multiple dimensions	UEQ, UEQ s, UEQ+, AttrakDiff, meCUE, meCUE2.0, iScale, UX Curve, MAX, UXSC, UXS, SUXES, MUX, sMUX, HED/UT, Bosch UX tool
Usability	SUS, CSUQ, QUIS, PSSUQ, MIPVA-U, NPS, SGUS, SEQ, USE
Flow	FSS, FKS, Flow4D16
Presence	ITC-SOPI, PQ, iPQ, SUS, EVEQ-GP
Emotion	SAM, 3E, 3E*, PrEmo, PANAS, I-PANAS-SF, Geneva Emotion Wheel, EmoCards, Emoticon, Emoti-SAM
Engagement	GEQ, UES, UES-SF
Workload	NASA-TLX, DALI, SMEQ, SEA
Game	CEGEQ, GEQ, EGameFlow, PGQ, GPQ, GUESS, PENS, MEEGA+
Simulator sickness	FMS, SSQ, MSAQ
AR-VR	VRNQ, UX in IVE, VRLEQ
Others	PPA, SASSI, SUISQ, SEI, VisAWI, INTUI, Reaction cards, Comfort scale, MIPVA-U, VOF, Fun toolkit, CAS, Godspeed questionnaire, IMI, PENS, UNeeQ, SGUS, UX needs scale.

[26] Slater-Usoh-Steed Questionnaire

Table 4.12 Most frequently used tools in UX studies covered by this SLR

Evaluation tool	Frequency
UEQ, UEQ-S, UEQ+	85
AttrakDiff	47
SUS	43
NASA-TLX	20
SAM	14
SSQ	9
meCUE	8
QUIS	8

Table 4.13 lists 43 UX evaluation tools that have been developed between 2010 and 2021. They are either case-specific or generally applicable tools. Some of them were developed from the ground up, while others were modifications of previous tools.

Table 4.13 Subjective UX evaluation tools identified in the SLR (2010-2021)

UX tool	Year	Field
User experience questionnaire for remote lab (UXQ4RL) **[121]**	2021	Education (Remote Lab)
User experience questionnaire for intraoperative video capture technology **[122]**	2021	Health (Interoperative camera)
Immersive virtual reality user experience questionnaire (IVRUX) **[123]**	2020	AR-VR
Virtual reality locomotion experience questionnaire (VRLEQ) **[124]**	2020	AR-VR
User Experience Questionnaire plus for voice assistants (UEQ+) **[96]**	2020	Others (Voice Assistant)
User experience framework for education games questionnaire (EDUGXQ) **[125]**	2020	Education (Educational game)
Questionnaire for technology-enhanced interaction with cultural heritage **[126]**	2020	Museum
Questionnaire for smart TV UX **[127]**	2019	Entertainment (Smart TV)
Formation, apprenant, systeme, enseignant relation learner experience (FASER LX) **[128]**	2019	Education
Instrument to evaluate user experience in virtual reality serious games **[129]**	2019	AR-VR (serious game)
HaBuT (happiness, burden, trust) instrument to measure user experience **[130]**	2019	Other (Content platform)
Questionnaire for the evaluation of serious games **[131]**	2019	Work (serious game)
Tangible kit to evaluate UX of older adults (Aestimo) **[132]**	2019	General tool (For elderly)
User experience questionnaire plus (UEQ+) **[50]**	2019	General tool
Mobile user experience (MUX), short form of mobile user experience (sMUX) **[133]**	2019	Others (Mobile devices)
Questionnaire for the evaluation of VR glasses system **[134]**	2019	AR-VR (VR glasses system)
Short user experience questionnaire **[135]**	2018	General tool
Modular evaluation of key components of user experience 2.0 **[94]**	2018	General tool
Questionnaire for the evaluation of social robot **[136]**	2018	Others (Social robot)
Questionnaire for the evaluation of Internet of things ecosystem **[137]**	2018	Others (IoT ecosystem)
Questionnaire for the evaluation of home appliances **[21]**	2018	Others (Home appliances)
Questionnaire for the evaluation of mobile application's interface **[138]**	2018	Others (App interface)
Questionnaire for the evaluation of mobile app **[139]**	2017	Education
User experience scorecard (UXSC), User experience scale (UXS) **[140]**	2017	General tool
Modular evaluation of key components of user experience (meCUE) **[8]**	2017	General tool
Quantified user experience (QUX) **[23]**	2016	Others (UX for organizations)
UX in Immersive virtual environments questionnaire (UX in IVE) **[141]**	2016	AR-VR

UX tool	Year	Field
Evaluating the user experience of interactive digital narrative **[142]**	2016	Entertainment (IDN)
Questionnaire for the evaluation of newspapers **[143]**	2016	Others (Digital newspaper)
Method for the assessment of experience (MAX) **[144]**	2015	General tool
Questionnaire for the evaluation of home appliances **[145]**	2014	Others (Home appliances)
Tactile user experience assessment board (TACTUX) **[146]**	2014	General tool
Interactive TV user experience questionnaire (ITVUX) **[147]**	2013	Entertainment (Interactive TV)
Questionnaire for the evaluation of mobile phone and services **[32]**	2013	Others (Mobile phones)
Chinese user experience questionnaire **[148]**	2013	General tool)
The user needs questionnaire (UNeeQ) **[149]**	2013	General tool
Questionnaire for the evaluation in an automotive context **[150]**	2013	Others (Car)
Questionnaire for the evaluation of location-based applications **[151]**	2013	Others (Location-based app)
Questionnaire for the evaluation of industrial robots **[152]**	2012	Work (Industrial robots)
iScale **[153]**	2012	General tool
UX Curve **[153]**	2011	General tool
Questionnaire for the evaluation of brain-computer interaction gaming **[154]**	2010	Entertainment (Gaming)

UX models, method, and frameworks have also been proposed in this period, building the foundation for the development of more UX evaluation tool (Table 4.14).

Table 4.14 UX model, methods, frameworks identified in the SLR (2010-2021)

UX model, method, framework	Year	Field
User experience evaluation index of augmented reality applications **[155]**	2020	AR-VR
User experience evaluation model of dynamic mathematics software **[156]**	2020	Education
User experience evaluation model of computer input devices **[157]**	2019	Others (Computer input device)
User experience evaluation model of smart jewelry **[158]**	2019	Others (Smar jewelry)
User experience evaluation method of 3D websites **[159]**	2018	Others (3D website)
User experience evaluation framework of e-moderation **[160]**	2015	Education (eModeration)
Empirical model of the user experience **[38]**	2014	Others (Headphone)
Assessment model of user experience of smart mobile phones **[161]**	2013	Others (Mobile phones)

4.3.11 UX dimensions

UX dimensions covered by evaluation tools and the studies we reviewed varied greatly, with some focused only on usability or task-oriented aspects, whereas others included hedonic and affective dimensions. Table 4.15 shows a grouping of UX dimensions in 13 main dimensions with some examples of components for each one. The number of times each dimension was present in the 325 studies is also indicated. The most frequent dimensions were usability, hedonic, utility, and sensory, which are all included in the UEQ and AttrakDiff questionnaires, the most frequently used tools. Surprisingly, the least frequent UX dimensions were the informational, social, physical, and cognitive dimensions. UX tools evaluating five or more dimensions are shown in Table 4.16. The complete list of UX tools and their dimensions can be found in Appendix B.

Table 4.15 UX dimensions and included sub-dimensions in the 325 studies covered by this SRL

UX dimension	Components of the dimension	No.	Percentage
Usability	Ease of use, learnability, controllability, error tolerance, efficiency, etc.	283	87.1
Hedonic	Stimulation identification, enjoyment, fun, novelty, desirability, etc.	271	83.4
Utility	Usefulness, helpfulness, stability, credible, practicality, reliability, etc.	222	68.3
Sensory (perceptual)	Visual (aesthetic), haptics, acoustics, beauty, etc.	191	58.8
Overall impression	Attractiveness, goodness, overall user experience	167	51.4
Judgment	Intention to use, recommending to others, endurability, loyalty, etc.	87	26.8
Affect	Positive emotion, negative emotion, valence, arousal	73	22.5
Engagement	Immersion, presence, flow, attention, absorption, involvement, etc.	70	21.5
Impact on body	Discomfort (comfort), workload, fatigue, sickness, headache, etc.	64	19.7
Informational	Content quality, content quantity, information structure, readability, etc.	39	12.0
Social	Social interaction, sociability, social acceptance, social actorship, etc.	34	10.5
Physical characteristic	Weight, shape, wearability, portability, delicacy, manageability, etc.	19	5.8
Cognitive	Learning, cognition effectiveness, learning performance, etc.	16	4.9
Others	Ownership, pride, competence, exploration, etc.	63	19.4

Table 4.16 UX tools with the most frequent dimensions evaluated

UX tool	Usability	Utility	Affect	Hedonic	Engagement	Social	Informational	Sensory	Impact on body	Judgment	Physical	Cognitive	Overall UX	Others	Total
MEEGA+	✓	-	-	✓	✓	✓	✓	✓	-	✓	-	✓	✓	✓	10
UX in IVE	✓	✓	✓	✓	✓	-	-	✓	✓	✓	-	-	-	✓	9
GUESS	✓	-	-	✓	✓	✓	-	✓	-	✓	-	-	-	✓	7
GEQ	-	-	✓	-	✓	✓	-	✓	✓	-	-	-	✓	✓	7
meCUE	✓	✓	✓	✓	-	-	-	✓	-	✓	-	-	✓	-	7
Park (2013)	✓	✓	-	✓	-	✓	✓	✓	-	-	✓	-	-	-	7
SASSI	✓	✓	✓	✓	-	-	-	-	✓	✓	-	-	-	-	6
UEQ+	✓	✓	-	✓	-	-	✓	✓	-	-	-	-	✓	-	6
EGameFlow	✓	-	-	-	✓	✓	-	-	✓	-	-	✓	-	✓	6
EDUGXQ	✓	✓	-	✓	✓	-	-	-	-	-	-	✓	-	✓	6
AttrakDiff	✓	✓	-	✓	-	-	-	✓	-	-	-	-	✓	-	5
CEGEQ	✓	-	-	✓	-	-	-	✓	-	✓	-	-	-	✓	5
QUIS	✓	-	-	✓	-	-	✓	✓	-	-	-	-	✓	-	5
UEQ	✓	✓	-	✓	-	-	-	✓	-	-	-	-	✓	-	5
UES	✓	-	-	✓	✓	-	-	✓	-	✓	-	-	-	-	5
UXQ4RL	✓	✓	-	✓	✓	-	-	-	-	-	-	✓	-	-	5
VRNQ	-	-	-	✓	✓	-	-	✓	✓	-	-	-	-	✓	5

4.4 Discussion

This SLR provided the state of the art on UX subjective evaluation studies published from 2010 to 2021. This discussion starts with a general overview of the products and user interactions evaluated in the studies reviewed. Then, we discuss the UX tools and methods that were used, followed by the UX dimensions that were found to be less frequently covered in UX tools. Finally, we highlight the importance of technology in improving UX evaluations and conclude with the main takeaways.

4.4.1 General overview

This SLR found that half of the UX studies were from Europe, and as well, out of the 43 new UX evaluation tools identified more than half (25) originated in Europe. Thus, Europe is the main contributor to UX studies with Germany and Finland being the most influential countries. In Asia, Indonesia had the highest number of papers, but they mainly used previously developed tools (72% used UEQ). China showed growth in this field by developing and applying various tools in different fields, and by providing the highest number of UX research studies on emerging technologies.

UX studies cover a wide spectrum of products, such as an industrial robotic arm, a fall assessment tool, a museum portal, a virtual dressing room, a classroom training simulator, games, and more. Consequent with the rising use of mobile devices globally, mobile applications had the highest number of UX evaluation studies. This trend can be expected to grow as handheld devices become more capable of running power-intensive applications and peripheral devices like AR-VR headsets, which currently require computers. Robinson and colleagues claimed that UX studies are falling behind emerging topics like AR/VR, gaming and mobile technologies [5]. This SLR found that this is no longer the case as, for instance, AR/VR accounted for 18.3% of all studies compared to 1.6% [5] and 2% [102] in previous SLRs.

UX evaluations were conducted on final products in 72.3% of the studies. Prototypes evaluations went from 10.3% in the SLR of Robinson et al. [5] to 25.9% (this SLR), with an increase in the number of studies from 2018 onwards. Compared to prototypes, final products provide the user with a complete set of features and characteristics that provides a comprehensive and multi-faceted UX. However, the researchers' concerns about a lack of UX studies in the early stages of product development remain [76]. However, “mixed prototype” or “augmented prototyping” has emerged

recently as a solution to simulate a final product using AR technology by projecting an artificial layer on the product physical model for usability testing purposes [162]. Moreover, using VR was also proposed to detect and correct design flaws faster than with physical prototypes and also to reduce their associated costs. However, touch and feel are lacking and have yet to be integrated to these new technologies [163].

All of the studies reviewed conducted UX evaluations after the interaction of the user with the product, and only 18.4% conducted an evaluation before and/or during the interaction as well. In the before-and-after evaluation studies, user expectations toward the expected use of the product were measured either by using common UX questionnaires like AttrakDiff [97] or specific UX questionnaires like SUXES (e.g., [164]) that specifically include expectation measurement. When interacting with children, fun toolkits were used to collect changes in their expectations before and after the interaction with a product, such as the Smileyometer, Again-Again table, Funoemeter, and Fun sorter tools. Just as Pettersson et al. [102] we found few studies taking into account expectation and its relation to UX, and as reported by Bargas-Avila and Hornback [16] we noted an overall lack of UX measurement before interaction.

The time that users spent with a product before UX evaluation was short (i.e., from a few minutes to a couple of hours) in 69.9% of the studies reviewed. Laboratory studies where users often have a limited time to interact with a product were more numerous than field studies.

4.4.2 UX tools and methods

Similar to Bargas-Avila and Hornbaek [16], no studies covering a product life cycle were found in this SLR. But longitudinal studies lasting up to 6 months to measure UX changes over time have been conducted throughout the last decade (e.g.,[165-168]). Some of the most famous long-term UX evaluation tools (UX curve [153] and iScale [169]) were developed in this period, as well as modified versions to address specific limitations of these original tools. For instance, MemoLine modified the UX curve to facilitate its use with children [170], while DrawUX was developed to collect UX data remotely [171]. UX graph was introduced as a revised version of UX curve, to capture in a single satisfaction graph, events occurring in different phases from the user's expectations before purchasing a product to the anticipation of future use [172]. Finally, the Experience Recollection Method (ERM) removed the exact timeline of the experience to reflect

the fact that episodic memory tends to fade with time, which in turn facilitated the use of the tool [173]. In addition to new tools, new approaches like the cross-sequential approach for data collection was presented to decrease the evaluation time by collecting UX data from different groups of users each being at a different phase of a product's usage [174].

The most common methods used for gathering qualitative data from users were interview and think aloud. Concurrent think aloud (CTA) is used during user interaction with a product, whereas retrospective think aloud (RTA) and interview are used after the interaction. The laddering technique and the valence method were used in UX studies like [175, 176] to improve the quality of interviews. The laddering technique starts the interview by talking about the product features and digs deeper into the subjective reasons behind the user's perceived importance of the features by repeatedly asking the "why" question [175]. In the valence method, users mark their positive and negative feelings during the use of a product and discuss those experiences with the researcher in a retrospective interview, enabling them to uncover underlying meanings and needs [176].

Among other qualitative methods, focus groups were used with experts for items development of new UX evaluation tools (e.g., [8]), while diaries were used to capture users' long-term experiences with the day reconstruction method (DRM) (e.g., [166]) or the experience sampling method (ESM) (e.g., [177]). Methods used in other domains found their way into the UX field and were tested to be used in conjunction with current UX evaluation tools and methods. For instance, Kujala and her colleagues [178] introduced the sentence completion method to the field of UX by conducting case studies comparing sentence completion to AttrakDiff. They found sentence completion to be less culturally biased compared to rating scales and easier to analyze compared to interviews. Similarly, co-discovery as a user testing method, in which two users interact with a product and freely talk about it, was used in the context of UX to elicit invaluable findings regarding user's perception, emotions and reactions [179].

This SLR identified all of the evaluation tools that were used in the UX studies as well as the new ones developed during 2010-2021. UX evaluation tools can be categorized either as general tools, that can be used with any product, or domain-specific tools that were developed for a particular domain. As opposed to Bargas-Avilla and Hornbaek [16] who reported that UX researchers would rather develop their own questionnaires than use the validated tools, we found that established

questionnaires for specific UX dimensions were being largely adopted since their use improves robustness of the findings and allows comparison of the results between studies.

Among the general tools, meCUE is based on the CUE model, and was first introduced in German [91] and then translated into English [92]. It has a modular structure comprising 1) product perceptions, 2) emotions, 3) consequences of use, and 4) overall UX modules. The evaluators select the modules that best apply to the UX dimensions they are interested in. meCUE is a validated tool that was compared with other well-known tools like AttrakDiff, UEQ, SAM and PANAS [8]. A major limitation of meCUE is that it is unsuitable for business-oriented applications such as an intranet, since users will disregard hedonic dimensions in these [93]. As a result, the second version of the tool (meCUE 2.0) splits the first module of product perceptions into two separate modules: 1) perception of instrumental product qualities and 2) perception of non-instrumental qualities [94]. Instrumental qualities include pragmatic and functional aspects of a product, while non-instrumental qualities cover hedonic aspects like beauty, status, and commitment [94].

The concept of modularity was embraced by the authors of the UEQ questionnaire and resulted in the development of UEQ+, a modular evaluation tool that provides a list of UX dimensions and corresponding evaluation items [50]. Modularity of this tool enables researchers to add more dimensions to UEQ+ depending on the product under study. For instance, Boos and Brau [95] added two haptic and acoustic UX dimensions for household devices, and Klein et al. [96] developed three new dimensions for the UX evaluation of voice assistants.

The belief that more domain-specific tools should be developed has many supporters on the grounds of the fact that standard evaluation tools cannot capture UX dimensions in all domains [93, 111, 147]. This led to the development of new tools for specific products like interactive TV [147], smart TV [127], home appliances [21, 145], and AR/VR [123, 141] to name a few. In addition to these, some current tools were tailored to be used specifically with new products or with new target populations. For instance, the Game Experience Questionnaire (GEQ) was modified to be used with BCI games [180], and the Chinese UX questionnaire was developed based on AttrakDiff with the assumption that UX dimensions that are commonly accepted in western cultures might not equally apply to eastern countries. Consequently, a tool with the following three dimensions, stimulation, pragmatic quality and conformity, was developed. “Conformity” was

defined as a dimension that shows how popular a product is among friends and how it is well-received by other users [148].

A number of UX evaluation tools use physical objects for the evaluation of UX, such as cards and boards. For instance, MAX captures the UX of users in four categories using cards [144], while Aestimo uses a tangible interface based on AttrakDiff specifically developed to capture the UX of elderly users. The interface uses tangible and familiar elements like buttons, knobs, and switches to be playful and at the same time increase the understandability of the evaluation process [132, 181]. Other tools like TACTUX use tactile properties to assess UX of a product [146].

The necessity of triangulating results from different tools and methods in UX evaluation has been established as a best practice in the field [102]. Consequently, the number of studies using multiple and varied tools has increased in the last decade, with 173 studies (55.4%) using two or more tools for their evaluation. Studies using a single tool that covers multiple dimensions (e.g., meCUE, UEQ, AttrakDiff) have also gained attention in recent years.

4.4.3 UX dimensions

This SLR identified more than 300 UX dimensions. We classified them under 13 main dimensions and determined which dimensions were addressed in different UX tools. Our findings uncovered drastic changes regarding the number of dimensions covered in UX studies compared to the results reported by Bargas-Avilla and Hornbaek [16]. In their review only 6% of studies evaluated four or more dimensions, whereas 78.8 % of the studies we reviewed covered at least four UX dimensions. Among the general UX tools, meCUE and UEQ+ cover the highest number of dimensions proving that recent UX tools are more comprehensive. As indicated earlier, modular tools enable researchers to modify them according to needs of their study and in the case of UEQ+, adding more dimensions is also possible. The multidimensionality of UX speaks to its flexibility and versatility, however it should be carefully managed as decomposing UX into several dimensions can lead to "dimensionality explosion," with a plethora of dimensions some of which may have no link to established constructs [16], while others measure the same thing but with different names. Recent studies like the one from Raptis et al. [182] examined the relationships between new factors (e.g., coolness) and other UX constructs to prevent further dimensionality increase.

Dimensions such as physical characteristics of products, cognitive, social, and informational have been less addressed in the UX studies we reviewed. The fact that software represented the majority of the products evaluated could explain why physical aspects of products were not frequent. Indeed, common UX tools such as AttrakDiff, meCUE and UEQ were originally developed for software products, thus do not include physical characteristics Some of the new tools like MUX and sMUX do take into account the role and impact of hardware on software in UX evaluation [133]. Raptis and colleagues [183] studied the effect of perceived hedonic quality of mobile devices on the UX evaluation of mobile applications, by comparing two different devices on factors like the style, visual design, age and the construction materials. Their results showed that pragmatic dimensions of mobile applications were affected by these factors [183]. Hence, physical characteristics not only impact directly the UX of physical products but also affect indirectly the UX of software products.

Another UX aspect that has received less attention in UX studies is the social dimension, which one could look at from three different perspectives. The degree to which a product enables the user to interact with others is one, while the acceptance of a user by others or by him/herself when using or owning a product is another one. The degree to which a product exhibits socially accepted behavior, which applies to robots or voice-based interactive products (e.g., [68]) is a third one. In this study we focused on the first two perspectives and categorized the UX dimensions accordingly. Among the general UX tools, UEQ and its extensions (i.e., UEQs, UEQ+) did not cover the social dimension, whereas AttrakDiff used only three bipolar items (isolating - connective), (separates me – brings me closer) and (alienating – integrating) mainly focused on the identification aspect of product usage. As for meCUE, it uses a "status" dimension to cover a user's self-image when using a product. The social dimension is better covered in other tools like the social module of GEQ[27] [65] that evaluates empathy with others, evoked feelings, and behavioral involvement caused by interaction with others. Collaboration and cooperation in social interactions are other facets that were included in recent studies [66, 67]. There is a need for UX tools or modules that

[27] Game Experience Questionnaire

better cover social dimensions in an era where new social media applications (e.g., Clubhouse, TikTok) and new features to existing products (e.g., bullet screen [184]) are announced daily.

The trustworthiness, quality, quantity, and structure of the information provided by the product are important to the overall UX that we categorized in the informational UX dimension. For instance, UEQ+ included trustworthiness and quality of content [50], the informativeness dimension was used with mobile devices [32], and trustworthiness and abundance of content were evaluated in the study of digital newspaper [143]. The informational dimension needs more attention now that we have both publishers generated (e.g., Netflix, Disney+, News websites, etc.) and user generated (e.g., YouTube, Instagram, TikTok, etc.) contents.

4.4.4 Technology and UX

Technology offers novel ways to collect user feedback when using the UX tools we reviewed. For instance, InteracDiff [185] implemented an interactive prototype that changes the presentation format of AttrakDiff to make it easier to understand and use for people outside the scientific community. Another example is the development of a pictorial scale called MAM[28] that enables users to remain in the virtual environment for the UX evaluation of AR-VR products [186]. Lastly, Meedin & Perera [187] developed a crowdsourcing platform for the evaluation of user interface design by gathering users and designers. Using an avatar, the platform asks questions to users and categorizes them in different personas. The designers can then evaluate their products with a target population through the intelligent user experience questionnaire (IUEQ) that is embedded in the platform, and thus generate distinct questionnaires for different types of users.

4.4.5 Limitations

SLRs have limitations that cannot be completely circumvented. In this SLR, the limitations include the selection of the search terms, researcher subjectivity in including or excluding papers and in categorizing sub-dimensions. We have incorporated four databases to ensure better coverage of

[28] Morph A Mood

UX studies and consulted a librarian specialist to use optimal search strings for each database. However, due to the language barrier, we limited the papers to only English publications.

4.4.6 SLR takeaways

This SLR shows that the tendency towards conducting UX evaluation earlier in the development phase (e.g., prototyping) has increased. This trend can address researchers' concern about a lack of UX studies in the early stages of product development. The case of AR/VR prototyping for hardware products is a good example of the application of UX evaluation in the early stages of development.

The low interaction time in UX studies using subjective and psychophysiological methods is another factor to be considered in UX evaluation. Short-term evaluations are less expensive and more feasible, but they are less reliable. Longitudinal studies where UX is measured over a long time are more difficult to organize and conduct but offer more reliable results about the actual usage of a product which provides higher quality feedback to ensure product success.

The variety of products and the contexts they were studied in shows how widespread UX has become. Similarly, the number of different evaluation tools used in UX studies has increased each year, rendering the evaluation more complicated particularly when many tools are developed to measure different dimensions of UX. UX evaluation requires experienced evaluators who can choose the best subset of tools to capture the many aspects of UX. As this SLR shows, modularity is the new trend in the development of UX evaluation tools (e.g., meCUE, UEQ+), as it increases the versatility of the tools to a great extent. Modular tools can be used in multiple contexts, giving the evaluator the option to customize the tool and to evaluate the set of UX dimensions most pertinent to the study, rather than using a combination of tools with different formats and rating scales that may overlap on some dimensions. From a practical point of view, using a same rating scale format simplify the assessment of different UX dimensions.

Modular tools covering a larger number of UX dimensions such as UEQ+ offer better access to the relevant dimensions associated with a product or a service. On the other hand, tools focusing on a limited number of dimensions tend to give a narrow view of the whole UX. This SLR found that the informational, social, physical, and cognitive dimensions appear to be less frequently covered

by current tools. For instance, the social dimension plays an important role in the evaluation of social products that have emerged in recent years and needs to be better considered in modular tools.

4.5 Conclusion

In this SLR, we reviewed the subjective UX evaluation tools and the dimensions they included for the period spanning from 2010 to 2021. Results showed the advent of modular tools in recent years and the advantages they bring through their potential to make the evaluation easier by gathering information on different dimensions with a same tool, and by being easy to customize to the particular product under study. We have also highlighted 4 groups of UX dimensions that were addressed less frequently in UX evaluation studies: informational, social, physical, and cognitive. Moreover, the current state of UX research regarding the countries with the most contributions to this field, the products evaluated, the evaluation time, the duration of product use before evaluation, the evaluation method, and the evaluation tools were presented to benefit researchers and practitioners alike.

CHAPTER 5 ARTICLE 2: IDENTIFYING THE IMPORTANCE OF UX DIMENSIONS FOR DIFFERENT SOFTWARE PRODUCT CATEGORIES

Ehsan Mortazavi, Philippe Doyon-Poulin, Daniel Imbeau, Jean-Marc Robert

This article was submitted to Interacting with Computers Journal on June 14, 2022

Abstract

Billions of users around the world use mobile applications and computer software to achieve their professional and personal goals. This situation drives User Experience (UX) researchers and practitioners to assess the importance of UX dimensions across different products, to facilitate the design, development, and evaluation of new products. To that end, this study surveyed a group of 200 end users and 8 UX experts from Canada to document the importance of 21 UX dimensions for 15 software product categories. The results confirmed that the importance of UX dimensions varies between product categories. Comparing the findings to those of similar studies conducted in Germany and Indonesia revealed that, while culture influences the rating of UX dimensions, the importance of UX dimensions is still determined by the product category. Comparisons between the importance ratings of UX dimensions between end users and experts and within end users were not significant in 77% and 97% of cases, respectively. Results showed that task-based product categories rely more on pragmatic dimensions (i.e., functionality and usability) while leisure-based products value hedonic dimensions (i.e., pleasure) as well. This study benefits researchers and practitioners by enabling them to select the most important UX dimensions for evaluating their products.

Keywords: User experience, UX dimension, UX evaluation, culture

5.1 Introduction

User Experience (UX) has been defined as “a multidimensional construct that defines the overall effect over time on the user of interacting with a system and service in a specific context”[45]. This definition highlights that UX is gained through the actual use of a product. It also emphasizes the trinity of users, products, and contexts in shaping UX. The influence of users on UX is through “users’ emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments” [4]. Gross and Bongartz [188] argued that UX is product-specific and the importance of UX dimensions varies between products. Moreover, new products, services, and technologies can affect UX by introducing new UX dimensions that have not been considered before [99]. Context of use can also impact the UX in different ways, such as physical, social, technical, and internal contexts when using a product [189].

UX dimensions that influence user-product interaction include, but are not limited to, pragmatic, hedonic, emotional, social, and physical dimensions. The pragmatic and hedonic dimensions, which respectively address functional and psychological aspects of UX, can be further divided into sub-dimensions. The more UX dimensions there are, the more time and effort participants will put into evaluation, and evaluators into analyzing the data. Some dimensions that are important for a product might be meaningless for another. Thus, knowing the important UX dimensions for different product categories helps keep UX evaluation focused and optimized. Depending on the product, one can evaluate only the relevant UX dimensions by using modular tools like meCUE 2.0 [94] or UEQ+ [50] or a combination of UX evaluation tools. For instance, Gross and Bongartz [188] used a combination of UX tools to evaluate three different products, and found clear differences between levels of importance of UX dimensions depending on the product type, such as goal-oriented vs. leisure-based. Santoso and Schrepp [99] compared the importance ratings of 16 UX dimensions for 15 software product categories and observed that differences in importance ratings exist at the category level. These studies pointed out the product specificity of UX. Additionally, we know that end users and UX experts alike can evaluate UX. It is interesting to know how these two groups rate the importance of UX dimensions for different product categories. In the case of similar judgments, one group could be chosen over the other based on the available time, cost, and goal of the evaluation. Jääskeläinen and Heikkinen [98] found that experts take

more UX aspects into consideration when evaluating a product compared to end users, including environmental and emotional aspects.

Users' cultural background should also be considered for UX evaluation, as people from different cultures have diverse perceptions, cognitions, and interaction styles with products [148]. Culture impacts the users' values, such as non-verbal communication [190], and products' design elements [191] [192]. Recent efforts have been made to incorporate cultural aspects into UX tools, such as the work of Li et al. that developed the Chinese UX questionnaire based on the AttrakDiff questionnaire [148]. They added a new dimension called "Conformity", which reflects the Chinese culture trait of giving value to others' opinions compared to one's own opinion in western societies. A major influence on the investigation of cultural aspects in UX studies is Hofstede's theory on cultural dimensions [190, 193], which defines culture as "the collective programming of the mind that distinguishes the members of one group or category of people from another". It includes six cultural dimensions, namely, power distance, individualism vs. collectivism, masculinity vs. femininity, uncertainty avoidance, long-term vs. short-term orientation, and indulgence vs. restraint [194]. Santoso and Schrepp [99] investigated the impact of culture on the importance ratings of UX dimensions for different software product categories in Germany and Indonesia. These two countries have distinct values on the six cultural dimensions of Hofstede's theory, making them suitable options for identifying the potential impact of culture on the importance of UX dimensions. Comparing the results of German [195] and Indonesian [196] students demonstrated that differences in importance of UX dimensions are caused more by product category than culture [99]. Therefore, we want to investigate to what extent these findings generalize to participants from another culture.

In this study, we replicated the work of Santoso and Schrepp [99] and extended it by adding five new UX dimensions, ratings from UX experts, and having participants from another country (i.e., Canada) as well. The results of a systematic literature review on the UX dimensions of interactive products revealed that social, informational and physical UX dimensions are underrepresented in current UX evaluation tools [197]. Thus, we included the social dimensions of sociability and social acceptance among the five new added dimensions. We chose Canada for the comparison with other cultures because it stands between Germany and Indonesia regarding the values of Hofstede's cultural dimensions. Therefore, this study has three objectives. First, to identify the important UX

dimensions for different software product categories. Second, to compare the importance rating of UX dimensions between end users and UX experts, and within the end-user group based on sociodemographic variables. Third, to perform a cross-cultural comparison.

This paper is organized in five sections. Following the introduction, section 2 presents the methodology for data collection and data analysis. Sections 3 and 4 show the results and a discussion in three subsections: importance of UX dimensions, comparisons between participants and cross-cultural comparisons. Section 5 presents a short conclusion.

5.2 Methodology

This study has been approved by Polytechnique Montreal's Research Ethics Committee (CER-2021-12-D) (Appendix C). Participants read and signed an informed consent form before taking part in the study.

5.2.1 Participants

Participants formed two groups: end users and UX experts. End users were from the general public, whereas UX experts were trained in the discipline and had years-long work experience. We collected data in two phases. The initial sample included 8 UX experts and 25 end users who all lived in Canada and were recruited from the personal contacts of the authors. End users received $20 per hour and UX experts received $50 per hour. A second sample of 175 end users was recruited through the SurveySwap website but received no monetary compensation. This sample of participants was also limited to Canada, using the website's custom requirements feature. Based on the feedback from the first sample of participants, two more product categories were added in the data collection for the second sample. The first sample of end users was required to answer all the product categories, whereas the second sample had the flexibility to answer none. We combined results regarding end users from both data collections into one sample. Table 5.1 presents the demographic data of the final sample, which shows how diverse the participants were in terms of age, gender, level of education, job status, familiarity with UX, and years of work experience in related fields for experts.

Table 5.1 Demographic data

Criteria		End user (N=200)	Expert(N=8)
Age	18-29	116	0
	30-39	57	2
	40-49	19	4
	50-59	7	1
	60-69	1	1
	70+	0	0
Gender	Male	93	3
	Female	105	5
	Other	2	0
Education	High school or equivalent	17	0
	Bachelor	86	0
	MSc.	80	4
	Ph.D.	17	2
	Other	0	2
Job status	Employed	75	6
	Unemployed	8	0
	Working and studying	29	1
	Student	88	0
	Retired	0	1
Familiarity with UX/HCI	Extremely familiar	19	8
	Very familiar	39	0
	Somewhat familiar	92	0
	Not so familiar	43	0
	Not at all familiar	7	0
Experience in the field of UX (Experts only)	Less than 3 years	-	0
	3-5 years	-	1
	5-10 years	-	0
	10-20 years	-	3
	+ 20 years	-	4

The two groups were well balanced on gender. Unlike experts, most end users were under the age of 39, and were students (59%) with a university-level education. Overall, 75% of end users were at least somewhat familiar with the fields of usability, UX, and HCI. Experts were chosen from both academia (3) and industry (5) to ensure that both viewpoints were considered in the evaluations, with work experience in visual design, cognitive ergonomics, UX research and UI development. Cross-cultural comparisons were made with the German and Indonesian samples of the work of Santoso and Schrepp [99] including 114 and 58 students, respectively.

5.2.2 Questionnaire

Data collection was done via an online questionnaire over a 2-month period for the first sample (September-October 2020), and over three months for the second sample (June-August 2021). It was designed on the SurveyMonkey website, and the initial participants accessed it through a link sent to them by email. For the second sample of end users, a list of product categories was made available for evaluation on the SurveySwap website. The questionnaire was written in English and designed based on the 16 UX dimensions from [99] with the addition of five UX dimensions (i.e., ease of use, error tolerance, sociability, social acceptance, and self-satisfaction; see below). Participants had to rate the importance of 21 UX dimensions on an 8-point Likert scale, ranging from meaningless (0) to extremely important (7). Participants were instructed to select Meaningless when they judged the UX dimension to be irrelevant for the product category. The first sample of end users and experts were asked to answer all product categories, whereas the second sample of end users was given the option to choose which product categories to evaluate. Moreover, two questions about the level of familiarity of participants with each product category and their frequency of use were also included in the survey. Participants had the option to write the UX dimensions that they felt were missing for each product category, as well as the name of the product in each category for which they filled the questionnaire.

In total, the survey contained 365 questions for the 15 product categories, and participants took on average 45 minutes to answer. The five UX dimensions added were chosen by three UX experts from a list of UX dimensions extracted from a systematic literature review performed earlier [197]. We decided to add ease of use as a new dimension. The description of perspicuity highlights the understandability and learnability of a product, so we wanted to investigate whether participants gave similar ratings for ease of use and perspicuity. Error tolerance was added as we expected this dimension to be important for productivity-related products like Word processing, Spreadsheet, and Online banking. The third dimension added was sociability, which is concerned with a product's ability to enable communication with other people. Social acceptance investigates how acceptable using a product is from the viewpoints of others or from the user's perspective. Finally, the dimension of self-satisfaction reflects the gratification a user gets when using a product. We expected this dimension to be seen in product categories involving learning or skill development. The list of 21 UX dimensions and their descriptions are presented in Table 5.2.

Table 5.2 UX dimensions and their descriptions (new dimensions indicated by *)

UX Dimension	Description
Beauty	The product is beautiful and attractive.
Ease of use*	It is easy to use the product.
Efficiency	I can achieve my goals with minimal time and physical effort. The product responds quickly to my input.
Error tolerance*	The product prevents me from making errors and helps me to recover from any that do occur.
Clarity	I find the user interface of the product looks tidy and clear.
Content Quality	The information provided by the product is always actual and of good quality.
Controllability	The product always reacts predictably and consistently to my input. I always have full control over the interaction.
Customization	I can adapt the product to my personal preferences or personal work style.
Identity	The product helps me to make contacts and to present myself positively.
Immersion	When I deal with the product, I forget the time. I completely sink into the interaction with the product.
Intuitive Usage	I can use the product directly without any learning or the help of other people.
Loyalty	Even if there are other equivalent products for the same tasks, I would not change the product.
Novelty	The design of the product is interesting and unusual. The original design catches my attention.
Perspicuity	It is easy to understand and learn how to use the product.
Self-satisfaction*	I feel satisfied with myself after using the product.
Sociability*	The product helps me to be sociable and connect with other people.
Social acceptance*	Using the product is socially accepted by others and my own norms.
Stimulation	I find the product stimulating and exciting. It is fun to deal with the product.
Trust	My data is in safe hands. The data will not be misused to harm me.
Usefulness	Using the product brings me advantages. It saves me time and effort and makes me more productive.
Value	I find the product makes a high-quality and professional impression.

We used the same 15 product categories as in Santoso and Schrepp's study for the comparisons. However, two new product categories were added in response to end users' feedback from the first sample. Moreover, we selected examples of each product category to be more easily understandable by Canadian users (Table 5.3).

Table 5.3 Product categories and their examples (new product categories are indicated by *). Products in parentheses moved to the new product categories for the second sample of the end-user

Product category	**Examples**
Word processing	Microsoft word, Google Docs, Apple's Pages, LaTeX.
Spreadsheet	Microsoft Excel, Google Sheets, Apple's Numbers.
Messenger	WhatsApp, Facebook Messenger, Snapchat, iMessage, Telegram.
Social network	Instagram, Facebook, (LinkedIn).
Video conferencing	Skype, Zoom, Face Time.
Web shop	Amazon, Walmart, Best Buy, eBay.
News portal	CBC News, CTV News, CNN, NBC news.
Booking system	TripAdvisor, Expedia, Trivago.
Info web page	Canada.ca, Montreal.ca, Polymtl.ca.
Learning platform	Udemy, Udacity, Lynda, SkillShare.
Programming tool	Eclipse, Visual Studio, Android studio.
Image processing	Adobe Photoshop, CorelDraw.
Online banking	NBC, RBC, Scotiabank.
Video portals	YouTube, (Netflix, Crave, Amazon prime video).
Games	PUBG, Fortnite, Dota2, League of legends.
Professional social Network*	LinkedIn, Meetup
Video Streaming*	Netflix, Crave, Amazon prime video.

Participants in the first sample of end users commented that some examples of product categories were confusing. For instance, having Instagram and LinkedIn in the same Social network product category made the rating exercise challenging for them. The nature of these two applications is different, as one is focused on entertainment, while the other is professional and work-related. The same applied to YouTube and Netflix in the Video portal category where the former allows users to express their thoughts in comments and discuss with others under each video, whereas the latter only allows like and dislike reactions to each video. This difficulty led us to add two more product categories to the survey of the second sample of end users, namely Professional social network and Video streaming services, in order to determine whether or not the differences were meaningful. This decision resulted in moving LinkedIn to the Professional social network category, keeping YouTube in the Video portal, and moving the rest into Video streaming.

We categorized the UX dimensions as in Table 5.4. The three UX dimensions of content quality, trust and loyalty did not have the same quality as the other dimensions included in the pragmatic and hedonic categories. Content quality is different from pragmatic and hedonic categories in that it is more concerned with the content a product provides rather than the product itself. Trust and

loyalty are UX dimensions that take shape after a period of working with a product. Furthermore, we grouped product categories into groups based on the contexts in which they are most commonly used (

Table 5.5).

Table 5.4 Categorization of UX dimensions (new dimensions indicated by *)

Category	**UX dimension**	
Pragmatic	Ease of use*	Customization
	Efficiency	Intuitive Usage
	Error tolerance*	Perspicuity
	Clarity	Usefulness
	Controllability	
Hedonic	Beauty	Sociability*
	Identity	Social acceptance*
	Immersion	Stimulation
	Novelty	Value
	Self-satisfaction*	
Other	Content Quality	
	Trust	
	Loyalty	

Table 5.5 Grouping of product categories (New product categories are indicated by *)

Group	**Product category**	
Work-related	Word processing	Image processing
	Spreadsheet	Professional social network*
	Programming too	
Personal task	Web shop	Online banking
	Booking system	
	Learning platform	
Entertainment	Social network	Video streaming*
	Video portal	
	Games	
Communication	Messenger	
	Video conferencing	
Information	News portal	
	Info web page	

5.2.3 Procedure

The participants read the instructions on how to fill out the questionnaire, and then proceeded to rate the importance of 21 UX dimensions for different product categories. To ensure a common understanding, a short description of each UX dimension was provided (Table 5.2), along with examples for each product category (Table 5.3). It was also mentioned that each category was not limited to those examples, and participants could give their judgment about other products belonging to the category. Instructions highlighted the fact that there was no correct or incorrect answer, and the answer should reflect the participant's personal assessment. The survey provided the option to go back and forth in the questionnaire or to stop at any time and resume later.

5.2.4 Data cleaning

In the first step, data was cleaned by removing the results of 39 participants who left the questionnaire incomplete, apparently answered randomly[29] or responded to the questionnaire in too short a time, i.e., under 25 minutes for answering all product categories for the first sample of end users and under 3 minutes for answering one product category for the second sample. The second step was to make sure that the participants actually used the product category they evaluated. To that end, for each product category, the results of the participants who indicated having "never" used a product, regardless of their familiarity with it, were removed. In addition, the results of those who indicated being "not so familiar" and "rarely" used a product category were eliminated. The final sample includes 200 end users and 8 UX experts. Table 5.6 shows the number of end users based on the frequency of use and familiarity with each product category.

[29] We used red-herring questions to find participants who randomly answered the survey.

Table 5.6 Number of end users' responses based on the frequency of use and familiarity per product category

Familiarity	Frequency of use	Word processing	Spreadsheet	Messenger	Social Network	Video Conferencing	Web shop	News Portal	Booking System	Info Web page	Learning Platform	Programming tool	Image Processing	Online Banking	Video Portal	Game	Professional SN	Video streaming
Extremely familiar	Always	16	6	20	23	8	8	1	2	0	1	5	0	11	26	4	1	16
	Usually	2	2	1	6	3	4	4	1	1	0	1	3	2	3	4	3	7
	Sometimes	0	1	0	0	1	2	2	2	0	0	0	0	0	1	1	2	0
	Rarely	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Very familiar	Always	12	6	7	5	2	0	5	0	0	0	2	0	6	8	0	1	2
	Usually	13	7	7	6	12	8	5	7	6	4	3	4	11	11	5	6	8
	Sometimes	3	2	1	1	5	7	2	5	4	7	1	4	0	4	1	4	2
	Rarely	0	1	0	0	0	0	1	1	1	2	1	1	0	0	0	0	0
Intermediate	Always	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	Usually	3	2	1	0	3	1	3	0	2	0	1	1	2	1	1	2	0
	Sometimes	2	11	3	2	5	9	13	10	8	11	6	10	6	5	9	8	3
	Rarely (4)	1	2	0	1	2	2	7	4	10	14	14	16	1	3	17	13	2
Total number of end users		53	40	40	44	41	41	44	32	32	39	34	39	39	62	42	40	41

5.2.5 Data analysis

In order to ensure that the newly added dimensions are not redundant we calculated Variance Inflation Factors (VIF) from the participants' ratings to assess multicollinearity among the 21 UX dimensions. Small VIF values indicate that the UX dimensions have low correlation with each other and thus are not redundant. Different thresholds of 10, 5, and 3 are provided for VIF to indicate multicollinearity between variables [198]. As in [199], we can use only the collinearity diagnostics feature of the linear regression dialog box of SPSS to calculate VIFs. The VIF values in this study were all less than 3 which is acceptable regardless of the selected cutoff value.

We used cluster analysis to categorize UX dimensions into important and unimportant for each product category. Among the clustering methods, we chose hierarchical clustering, which is suitable for a small amount of data and represents clusters in a tree-like visual graph called a Dendrogram. Using an agglomerative approach, each dimension is in a separate cluster at first. After each clustering step, dimensions with more similarities are grouped together until the end, when all dimensions are collected into one cluster [200]. We used Ward's method, which is similar to k-means clustering in that it minimizes the variances of members within a cluster. One characteristic of Ward's method is that it tends to produce even-sized clusters, which in the case of our study helps to have both clusters with more representative dimensions [200].

We used Multi-Dimensional Scaling (MDS) to show which product categories have similar UX dimension ratings. This visual technique puts product categories with comparable UX dimension ratings in closer proximity. Following the method used in [99], we used Pearson correlation of importance ratings as the similarity matrix for MDS to show the distance between products. We also used MDS to group UX dimensions into two groups of hedonic and pragmatic dimensions.

According to the Central limit theorem, we can assume normal distribution for sample sizes over thirty and use parametric tests [199]. Although there are at least 30 end users per product category, the total number of UX experts is only 8. Therefore, we used the Mann-Whitney U test, the non-parametric equivalent of the independent t-test, to compare the ratings of UX dimensions between end users and experts. To investigate the differences within the end users' group, we repeated the Mann-Whitney U test based 1) on users' frequency of use and familiarity with the products and 2) on gender, for all 17 product categories. The first group included users who were extremely or very familiar with a product category and always or usually used it and the rest belonged to second group (see Table 5.6).

Comparing our results with those of Santoso and Schrepp [99] required calculating the mean importance rating for each UX dimension of each product category that was featured in both studies. To that end, product-moment correlation of the mean importance rating of each product category between end-user samples of three countries was measured. We carried out the same procedure on the ranking values between samples for the comparison of UX dimensions' rankings per product category. Moreover, the average importance rating of UX dimensions for all product categories and the average importance rating of hedonic and pragmatic dimensions for the five groups of categories were measured to show the similarity and differences between the three cultures. Data analysis was performed on IBM SPSS statistics 27.

5.3 Results

Results are presented in three sections in accordance with the three objectives of the study. In the first section we identify the important and unimportant UX dimensions for each product category and investigate these dimensions for the two newly added product categories. In the second section we compare the ratings of UX dimensions between end users and experts. In the third section, we present cross-cultural comparisons between our findings and those of Santoso and Schrepp [99].

5.3.1 Important UX dimensions for different product categories

After having added the five new dimensions, VIF values for all UX dimensions ranged from 1.15 to 2.13. Because they are all under the threshold of 5, they indicate low multicollinearity among the dimensions. Results of the hierarchical cluster analysis are shown in Table 5.7 for the end users' data. Ease of use, intuitive usage, and perspicuity were the UX dimensions that were important for all product categories, whereas novelty, beauty, and identity were only important for one or two product categories. Professional social network (SN in Table 5.7) was the category with the most UX dimensions rated as important (16) whereas Messenger had the smallest number of important UX dimensions (6). Among all product categories, Game had most of the hedonic dimensions categorized as important. In addition to clustering, UX dimensions with the highest rankings per product category can be used as the most important dimensions as presented in Appendix D.

Table 5.7 Clustering of UX dimensions into important (checked) and unimportant (minus) per product category based on end- users' data

UX dimension (abbreviation)	Messenger	Info Web page	Booking System	News Portal	Web shop	Word processing	Spreadsheet	Image Processing	Programming tool	Online Banking	Learning Platform	Video streaming	Video Portal	Professional SN	Social Network	Video Conferencing	Game
Content quality (CQL)	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓
Customization (CUS)	-	-	-	-	-	✓	✓	✓	✓	✓	-	✓	✓	✓	-	✓	-
Perspicuity (PER)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Efficiency (EFF)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
Immersion (IMM)	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-	✓
Intuitive usage (INT)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Usefulness (USF)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
Novelty (NOV)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓
Beauty (BEA)	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	✓
Identity (IDN)	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-
Controllability (CON)	-	-	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓
Stimulation (STM)	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-	✓
Clarity (CLR)	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Loyalty (LOY)	-	-	-	-	-	-	-	-	✓	✓	-	-	-	-	-	-	-
Trust (TRS)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	-
Value (VAL)	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	-
Ease of use (EOU)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Error tolerance (ERR)	-	-	✓	-	✓	✓	✓	✓	✓	✓	-	-	✓	✓	-	✓	✓
Sociability (SOC)	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	-
Social acceptance (SOA)	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	-
Self-satisfaction (SLF)	-	-	-	-	-	-	✓	✓	✓	✓	✓	✓	-	✓	✓	-	✓

Product categories with comparable importance ratings of UX dimensions can be identified in Table 5.7, however MDS's visual representation is easier to interpret (Figure 5.1). Product categories having similar ratings among all UX dimensions are located closer to each other (see Figure 5.1-left). Task-focused products, both professional and personal, informational, and video service are all placed near each other. Although not closely grouped together, product categories of a social nature are placed in distance from others (Triangles). Similarly, the Game category, being focused on hedonic dimensions, is located far from the rest.

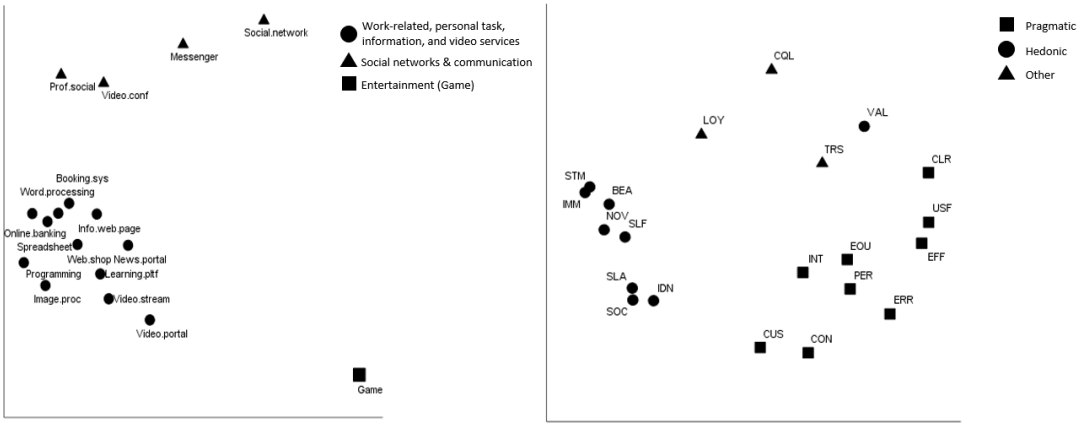

Figure 5.1 multi-dimensional scaling of product categories (left), UX dimensions (right)

The results of the MDS analysis on the UX dimensions are shown in the right half of Figure 5.1, with hedonic dimensions on the left (circles) and pragmatic dimensions (squares) on the right. The only exception was value (VAL) that was originally classified as a hedonic dimension, but it appeared closer to pragmatic dimensions. Moreover, loyalty, trust, and content quality (triangles) showed up in between the two groups. Results also show that social acceptance, sociability, and identity can be seen as a separate category.

Mean importance ratings and rankings of UX dimensions for the two social network categories show that users put more emphasis (i.e., more than twice in ranking) on usefulness, efficiency, value, sociability and customization of the Professional social networks compared to the Social network (see Table 5.8). The Pearson correlation of the ranking value of Social network and Professional social network was low at 0.455, but statistically significant ($p = .038$).

Table 5.8 Ranking and mean importance rating of UX dimensions for social network and professional social network categories (ranking from 1 to 21 and rating from 0 to 7)

Product category		Content quality	Customization	Perspicuity	Efficiency	Immersion	Intuitive usage	Usefulness	Novelty	Beauty	Identity	Controllability	Stimulation	Clarity	Loyalty	Trust	Value	Ease of use	Error Tolerance	Sociability	Social acceptance	Self-satisfaction
Social network	Ranking	4	12	7	13	16	5	17	21	15	8	9	14	10	19	3	18	2	20	1	6	11
	Rating	6.0	5.4	5.8	5.3	5.1	5.9	5.0	4.7	5.2	5.7	5.7	5.3	5.5	5.0	6.2	5.0	6.3	4.7	6.5	5.9	5.5
Professional social network	Ranking	4	6	11	3	21	9	1	20	19	13	12	17	10	18	7	5	2	16	8	14	15
	Rating	5.8	5.7	5.5	5.9	3.3	5.6	6.3	3.4	3.5	5.4	5.5	3.7	5.6	3.7	5.7	5.7	5.9	5.2	5.6	5.4	5.3

Video portal and Video streaming categories did not show much of a difference, except the ranking of the value, stimulation, and efficiency dimensions, which were noticeably different (i.e., more than twice) (see Table 5.9). The ranking values of UX dimensions for these two categories had a strong correlation of 0.826, which is highly significant ($p = .000004$).

Table 5.9 Ranking and mean importance rating of UX dimensions for video portal and video streaming categories (ranking from 1 to 21 and rating from 0 to 7)

Product category		Content quality	Customization	Perspicuity	Efficiency	Immersion	Intuitive usage	Usefulness	Novelty	Beauty	Identity	Controllability	Stimulation	Clarity	Loyalty	Trust	Value	Ease of use	Error Tolerance	Sociability	Social acceptance	Self-satisfaction
Video portal	Ranking	1	7	6	8	9	3	14	16	15	21	10	4	5	18	11	13	2	17	20	19	12
	Rating	6.4	5.5	5.5	5.5	5.4	6.0	5.0	4.4	4.9	3.1	5.3	5.6	5.5	4.1	5.1	5.1	6.1	4.3	3.1	4.0	5.1
Video streaming	Ranking	1	13	6	3	12	4	8	17	16	20	7	11	9	18	10	5	2	14	21	19	15
	Rating	6.6	4.8	5.6	6.0	5.2	5.9	5.6	3.5	3.9	2.4	5.6	5.4	5.6	3.5	5.5	5.8	6.3	4.7	2.3	2.9	4.7

5.3.2 Comparison of end users and experts' ratings of UX dimensions

Table 5.10 reports the results on the comparisons of end users' ratings with those of the experts for the 21 UX dimensions and the original 15 product categories, except for the Programming tool and Game categories due to a lack of expert data. The p-values for the Mann-Whitney U tests are shown in Table 5.10, with bold values denoting dimensions where the two groups differ significantly. Results show no significant difference between end users' and experts' ratings for the five UX dimensions of identity, loyalty, value, sociability, and social acceptance, across all product categories. Spreadsheet and Social network had the lowest number of differences between two

groups. The mean ranks of the experts were higher than those of end users for all 62 occurrences where a significant difference was found.

Table 5.10 P-values of the Mann-Whitney U tests between end users' and experts' ratings

UX dimension	Word processing	Spreadsheet	Messenger	Social Network	Video Conferencing	Web shop	News Portal	Booking System	Info Web page	Learning Platform	Image Processing	Online Banking	Video Portal
Content quality	0.278	0.483	0.692	0.274	0.105	**0.014**	0.082	**0.043**	0.151	0.194	0.163	**0.030**	0.197
Customization	0.251	0.074	0.625	0.990	0.200	0.433	0.969	0.053	0.695	**0.042**	**0.024**	0.287	0.687
Perspicuity	0.107	0.156	0.243	0.659	**0.027**	0.560	0.213	**0.042**	0.094	**0.004**	**0.005**	0.153	0.126
Efficiency	**0.044**	0.128	0.208	0.103	**0.006**	0.122	0.624	0.381	0.827	0.068	**0.006**	0.295	0.056
Immersion	0.444	0.191	**0.039**	0.137	0.485	0.891	**0.011**	0.433	0.393	0.307	0.060	0.399	**0.006**
Intuitive usage	0.106	0.184	0.120	0.055	**0.003**	0.089	**0.028**	**0.031**	**0.030**	**0.006**	**0.006**	**0.032**	**0.038**
Usefulness	0.045	0.089	0.351	**0.022**	**0.029**	0.089	0.533	0.362	0.753	0.079	**0.023**	0.389	0.400
Novelty	0.957	0.942	0.054	0.103	0.104	0.389	**0.001**	0.060	**0.0002**	0.145	0.989	**0.043**	0.122
Beauty	0.307	0.431	0.092	0.373	0.100	0.484	**0.004**	0.062	**0.015**	**0.038**	0.143	**0.011**	0.071
Identity	0.468	0.330	0.112	0.406	0.553	0.691	0.867	0.473	0.810	0.652	0.886	0.374	0.723
Controllability	**0.012**	**0.008**	**0.002**	**0.031**	**0.018**	**0.044**	0.072	**0.032**	0.352	**0.001**	**0.004**	0.119	**0.006**
Stimulation	0.802	0.519	0.160	0.774	0.130	0.387	0.078	0.644	0.280	0.056	0.428	0.509	**0.048**
Clarity	**0.016**	**0.034**	**0.039**	0.085	**0.002**	**0.005**	0.052	0.147	0.407	**0.003**	**0.019**	0.189	**0.022**
Loyalty	0.738	0.831	0.319	0.478	0.785	0.814	0.727	0.192	0.183	0.938	0.277	0.171	0.808
Trust	0.073	0.429	0.856	0.773	0.084	**0.037**	0.590	0.120	0.734	0.080	0.326	0.166	0.221
Value	0.092	0.429	0.856	0.773	0.084	**0.037**	0.590	0.120	0.734	0.080	0.326	0.060	0.308
Ease of use	0.167	0.292	0.766	0.722	0.095	0.790	0.335	0.404	0.616	**0.005**	**0.042**	0.096	**0.014**
Error tolerance	**0.021**	0.094	0.120	0.345	0.267	0.394	0.432	0.446	0.986	0.206	0.195	0.725	0.743
Sociability	0.753	0.490	0.190	0.920	0.772	0.621	0.887	1	0.251	0.220	0.826	0.650	0.940
Social acceptance	0.073	0.352	0.917	0.774	0.723	0.795	0.391	0.430	0.609	0.475	0.954	0.117	0.815
Self-satisfaction	**0.007**	0.085	0.143	0.160	0.377	0.149	0.624	0.049	0.292	0.062	**0.003**	**0.034**	0.427

The Pragmatic dimensions amounted for 43 out of the 62 instances where a significant difference was found, with controllability, clarity, and intuitive usage being the dimensions with the highest number of differences between the two groups of participants.

Results of the comparisons within end users with different frequency of use and familiarity with the products showed rather close ratings, with only 15 pairs (out of 357) showing significant differences. The first group (i.e., more familiar) gave higher mean rank for 11 pairs including both pragmatic and hedonic dimensions. Among product categories, Spreadsheet had the greatest number of differences for the pragmatic dimensions of customization, efficiency, usefulness, and error tolerance rated higher by those who were more familiar with it.

Regarding the gender differences, we only found 12 instances (out of 357) where a significant difference was found between two groups. The highest number of differences was found in Social network product category with perspicuity, immersion, and intuitive usage as dimensions with

higher mean ranks for females. Error tolerance for Word processing, usefulness for Image processing, and clarity and sociability for Online banking were the only four dimensions with higher mean ranks for males.

5.3.3 Cross-cultural comparison

This section presents the results of the Canadian end users' evaluation on the original set of 16 UX dimensions and 15 product categories as in Santoso and Schrepp's study [99], whose participants were from Germany and Indonesia. Appendix E provides the ranking and the mean importance rating of each UX dimension for each product category.

Table 5.11 and Table 5.12 show the Pearson correlation between samples for the mean importance rating and the ranking of UX dimensions, respectively. The last column of tables 11-12 shows the correlation of the average importance rating of UX dimensions and the average ranking of UX dimensions for all product categories. Results showed moderate to very high correlations between samples in both tables. We found the highest correlations for Word processing, Spreadsheet, Booking system, Learning platform, Programming tool, Image processing, and Online banking between Canadians and the other two samples. The only exception was the Video portal category, with its correlation not being statistically significant between Canadians and Indonesians. Still moderate, we observed the lowest correlation for the Social network category between Canadians and Indonesians. Overall, correlations were higher between Germans and Canadians for both rankings and ratings.

Table 5.11 Pearson correlation of the mean importance ratings for each product category between samples

Country*	Word processing	Spreadsheet	Messenger	Social Network	Video Conferencing	Web shop	News Portal	Booking System	Info Web page	Learning Platform	Programming tool	Image Processing	Online Banking	Video Portal	Game	Average
CAD-IND	0.97	0.96	0.72	0.56	0.77	0.82	0.67	0.85	0.81	0.89	0.95	0.92	0.86	0.44	0.89	0.89
CAD-GER	0.95	0.95	0.80	0.72	0.88	0.82	0.72	0.94	0.81	0.94	0.92	0.87	0.92	0.62	0.86	0.94
GER-IND	0.94	0.95	0.76	0.65	0.83	0.91	0.91	0.89	0.93	0.93	0.92	0.91	0.94	0.84	0.89	0.93

*CAD: Canada; IND: Indonesia: GER: Germany

Table 5.12 Pearson correlation of the UX dimensions ranking for each product category between samples

Country*	Word processing	Spreadsheet	Messenger	Social Network	Video Conferencing	Web shop	News Portal	Booking System	Info Web page	Learning Platform	Programming tool	Image Processing	Online Banking	Video Portal	Game	Average
CAD-IND	0.97	0.93	0.77	0.54	0.76	0.85	0.60	0.93	0.79	0.91	0.92	0.89	0.87	0.54	0.80	0.87
CAD-GER	0.93	0.95	0.75	0.78	0.88	0.80	0.66	0.89	0.79	0.92	0.96	0.88	0.80	0.70	0.79	0.93
GER-IND	0.91	0.94	0.71	0.59	0.81	0.86	0.85	0.89	0.84	0.93	0.90	0.85	0.76	0.83	0.82	0.90

*CAD: Canada; IND: Indonesia: GER: Germany

Figure 5.2 shows the results of the average importance rating of hedonic and pragmatic dimensions for each group of products. Regarding the hedonic dimensions, the three cultural groups gave on average the highest ratings for the Entertainment products and the lowest for Work-related products (Figure 5.2– left). Indonesians consistently rated hedonic aspects the highest among all three cultures. Regarding the pragmatic dimensions, the highest mean importance ratings were associated with the Work-related products, whereas the lowest means were for Information products, in all three countries (Figure 5.2-right).

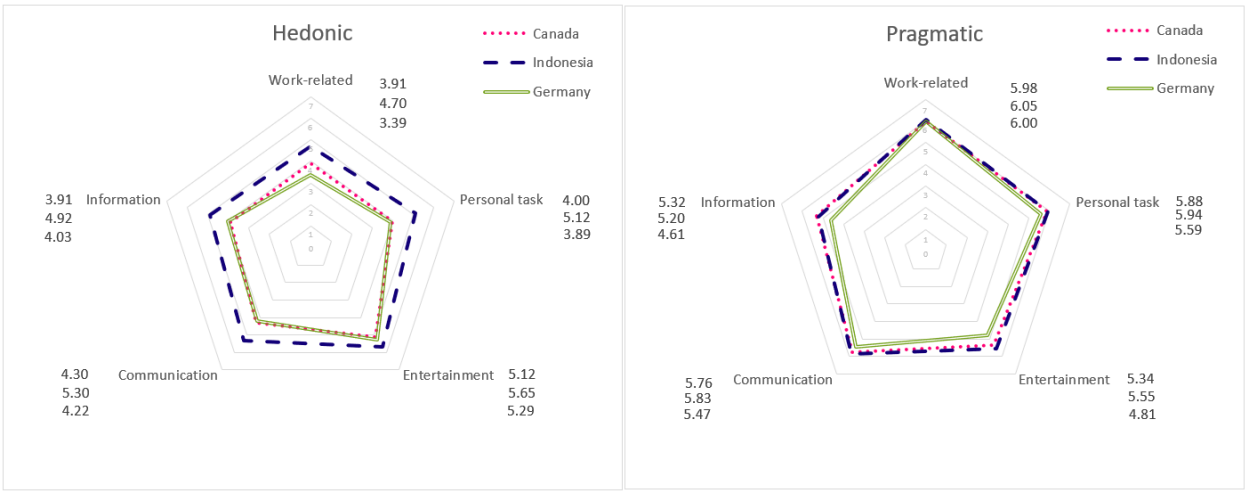

Figure 5.2 Average importance rating of hedonic and pragmatic dimensions for each group of products (ratings 0-7)

5.4 Discussion

In this section, we first identify the important UX dimensions for different product categories, then discuss the comparison of the UX ratings among participants and finally explore cross-cultural comparisons of the UX ratings.

5.4.1 Important UX dimensions

Identifying which UX dimensions are important for different product categories allows designers to improve the functions and features of their products. Focusing on the most important UX dimensions for a product category also facilitates evaluation by avoiding lengthy and costly evaluation sessions that measure irrelevant UX dimensions.

Our findings (Table 5.7) revealed that certain UX dimensions, such as ease of use, intuitive usage, and perspicuity, were important regardless of product category. Other dimensions like beauty, novelty, and value were important only for one or two product categories. These results are in accordance with Jordan's [201] hierarchy of consumer needs, which begins with functionality and proceeds to usability and pleasure at higher levels. Based on our results, we can argue that task-based product categories rely more on pragmatic dimensions (i.e., functionality and usability) while leisure-based products value hedonic dimensions (i.e., pleasure) as well.

Product categories associated with Work, Personal tasks, and Information groups (see Figure 5.1) shared more UX dimensions except for the Professional social networks (see Table 5.13). We observed that newly added dimensions had an impact on better differentiating the product categories. For instance, error tolerance was important for task-based categories for which the possibility and cost of making errors were higher, that is, those that dealt with large amount of input by users (e.g., Word processing, Spreadsheet, Programming tool) and those that involved money transactions (e.g., Web shop, Booking system, and Online banking). Findings also showed that self-satisfaction is an important UX dimension for the product categories whose outcome involved learning (e.g., Learning platform), development (e.g., Programming tool, Image processing), or enjoyment (e.g., Game, Video streaming).

The two groups of Entertainment and Communication (Figure 5.1-left), in which hedonic dimensions were important, were noticeably different from the other three groups, with Game

standing out from the rest. The results of MDS for UX dimensions showed that identity, sociability, and social acceptance were grouped together (Figure 5.1-right). The same social dimensions distinguished Social networks and Video conferencing from other product categories. These dimensions have the potential to be evaluated separately from the hedonic dimensions, which is consistent with what [197] stated regarding the need for increased attention to social dimensions in UX studies.

The UX dimensions identified as important for each product category (as in Table 5.13) can be measured either by using a combination of UX evaluation tools or a modular evaluation tool. UEQ+ [50] is an example of a modular framework that enables evaluators to select the most relevant UX dimensions for the evaluation of their products. Other studies [95, 96] have added more dimensions to UEQ+, demonstrating its potential to entail more dimensions and to be used for a wide range of products.

Table 5.13 Important UX dimensions for different software product categories

Product category	Important dimensions
Word processing	ease of use, efficiency, error tolerance, clarity, controllability, customization, intuitive usage, perspicuity, trust, usefulness, value
Spreadsheet	ease of use, efficiency, error tolerance, clarity, controllability, customization, intuitive usage, perspicuity, trust, usefulness, value
Messenger	ease of use, efficiency, clarity, controllability, immersion, intuitive usage, perspicuity, trust, usefulness
Social network	ease of use, efficiency, error tolerance, clarity, content quality, controllability, identity, intuitive usage, perspicuity, self-satisfaction, sociability, social acceptance, trust, usefulness
Video conferencing	ease of use, efficiency, error tolerance, clarity, controllability, customization, intuitive usage, perspicuity, sociability, social acceptance, trust, usefulness, value
Web shop	beauty, ease of use, efficiency, error tolerance, clarity, content quality, controllability, intuitive usage, perspicuity, trust, usefulness, value
News portal	beauty, ease of use, efficiency, clarity, content quality, controllability, immersion, intuitive usage, novelty, perspicuity, trust, usefulness, value
Booking system	ease of use, efficiency, error tolerance, clarity, content quality, controllability, intuitive usage, perspicuity, trust, usefulness
Info web page	beauty, ease of use, efficiency, clarity, content quality, intuitive usage, perspicuity, trust, usefulness
Learning platform	beauty, ease of use, efficiency, clarity, content quality, controllability, intuitive usage, perspicuity, self-satisfaction, trust, usefulness, value
Programming tool	ease of use, efficiency, error tolerance, clarity, controllability, customization, intuitive usage, loyalty, perspicuity, self-satisfaction, trust, usefulness, value
Image processing	ease of use, efficiency, error tolerance, clarity, controllability, customization, intuitive usage, perspicuity, self-satisfaction, trust, usefulness, value

Product category	Important dimensions
Online banking	ease of use, efficiency, error tolerance, clarity, controllability, customization, intuitive usage, loyalty, perspicuity, trust, usefulness, value
Video portals	ease of use, efficiency, error tolerance, clarity, controllability, customization, immersion, intuitive usage, perspicuity, stimulation, trust, usefulness, value
Games	beauty, ease of use, error tolerance, clarity, controllability, immersion, intuitive usage, novelty, perspicuity, self-satisfaction, stimulation
Professional social Network	ease of use, efficiency, error tolerance, clarity, content quality, controllability, customization, identity, intuitive usage, perspicuity, self-satisfaction, sociability, social acceptance, trust, usefulness

The importance of UX dimensions can be influenced by factors related to the context, users, and their activities. For instance, loyalty to a product can be mandatory, like using the banking app in which you have an account, or optional, like a programmer who uses a specific programming language. It can vary based on the user's skills, like pro-users who use shortcuts and customize their working environment in an image processing or spreadsheet software. It can also differ depending on the target user, for instance, a person whose job is content creation on social networks or video portals might give a different importance rating of UX dimensions compared to someone who uses these products for entertainment purposes only. In other words, although having a general overview of important dimensions for each product category is useful, one should also consider the subtle hints that differentiate products. UX dimensions of products within the same product category, for example, can be rated differently should their purpose differ. To that end, we compared the rating and ranking of UX dimensions between entertainment-focused Social networks and business-oriented Professional social network as well as Video portals and Video streaming services. Our findings suggest that separating social networks into two categories allows better detection of differences. However, there was no need to have two separate categories for Video portals and Video streaming services because their ratings and rankings on UX dimensions were comparable.

5.4.2 Comparison between participants

Our second goal was to see how the two groups of end users and experts rated UX dimensions per product category. In 77% of cases, there was no significant difference between the ratings of the two groups (211 out of 273). For the rest, experts gave higher ratings for all UX dimensions compared to end users. We found that all the pragmatic dimensions that were significantly different between end users and experts (see Table 5.10) were clustered as important by the users (see Table

5.7), except for controllability and clarity for Messenger and customization for Learning platform categories. This finding shows that regardless of a significant difference, pragmatic dimensions were important to both groups.

Results showed that end users were more concerned with pragmatic rather than hedonic dimensions when it comes to task-based products. On the other hand, experts gave higher ratings for hedonic dimensions such as beauty and novelty for product categories like News portal, Information web page, Learning platform and Online banking. From the experts' perspective, these task-based products can benefit from the hedonic dimensions to be more inviting and attractive to the users. Our findings are in line with those of [98], demonstrating that experts take more UX aspects into consideration when asked to evaluate a product.

Comparisons within the end-user group yielded interesting results. Our findings showed significant differences only for 12 of the comparisons (out of 357) between genders. Similarly, 15 pairs (out of 357) were significantly different between the end users based on frequency of use and familiarity with the product category. We can argue that these two factors did not make a meaningful difference in the rating of the UX dimensions. However, the less familiar users are with a product category, the more likely they are to evaluate their expectations of a product rather than its actual use. For instance, users the least familiar with Spreadsheets rated customization, efficiency, usefulness, and error tolerance significantly lower, which is unexpected for a task-based product category.

5.4.3 Comparison between cultures

The last goal of this study was to investigate to what extent culture affects the evaluation of the importance of UX dimensions. The mean importance ratings of the UX dimensions were different between the three cultures. According to Figure 5.2, the Indonesians gave higher ratings for all UX dimensions, specifically hedonic dimensions, compared to the other two cultures. The findings of [192] on the UX evaluation of smartphones based on Hofstede's cultural dimensions showed that users from high power distance cultures, such as China and India, are cautious in freely expressing their opinions and giving negative feedback. Moreover, cultural impact on website development showed that cultures with higher values on Hofstede's theory feminine dimension emphasize on the hedonic values such as website aesthetics [202]. Thus, the differences in ratings of the

Indonesians can be traced back to their culture, which has higher power distance and lower masculinity values compared to German and Canadian cultures [203]. However, regardless of differences in mean importance values, the high correlation between the mean importance rating and ranking of UX dimensions (Table 5.11 and Table 5.12) proved that participants from all three countries had a similar judgment of the importance of UX dimensions for each product category. The only two product categories with a lower correlation were Social network and Video portals. As discussed in the previous section, depending on what products have been evaluated we could see differences in importance ratings most notably for the Social network product category.

5.4.4 Limitation

The limitations of our work are as follows: first, ratings were not done for a specific product per category. Although covering different products enabled us to measure the overall judgment of participants on a product category, differences between products in each group can impact the results. Second, we did not instruct participants to consider a specific device when evaluating a product category. The UX of an online word document (e.g., Google Docs®) can be different when using a computer with a cursor and mouse compared to a mobile phone. Third, based on our sample size, we could not investigate the impact of other factors like age on ratings since most of the participants were young. Fourth, since UX is flexible and evolves over time the list of important UX dimensions is not static and requires to be reviewed and updated over time. Last, we had to rely on the correlations between the rankings and the mean importance ratings of UX dimensions between different cultures because we did not have access to the raw data of Santoso and Schrepp's study.

5.4.5 Take-aways

Based on our findings, the main take-aways are as follows:

- The importance of UX dimensions varies depending on the product category.
- UX experts' important UX dimensions for different product categories overlapped with those of end users yet included more hedonic dimensions.

- The importance of UX dimensions for evaluating different categories of products was similar across different cultures. However, culture can affect the rating values for different UX dimensions.
- UX designers can use the list of important UX dimensions for each category of products to prioritize dimensions that will have the greatest impact.

5.5 Conclusion

This study investigated the importance of UX dimensions for different product categories and compared the evaluation results of 200 participants in Canada to those of 114 Germans and 58 Indonesians. We confirmed the conclusion of prior research that the product category had a greater impact on UX evaluation than cultural differences. Additionally, in most cases, there was no significant difference in the evaluation of the importance of UX dimensions when factors such as familiarity with UX (i.e., end user vs. experts), frequency of use of a product, familiarity with the product, and gender were investigated. This study highlighted the importance of five new UX dimensions, most notably sociability and social acceptance. The lack of attention to these dimensions in UX evaluation tools requires more studies to cover these social aspects. We are currently in the process of developing a tool for the evaluation of the social dimensions of UX.

CHAPTER 6 ARTICLE 3: DEVELOPMENT AND VALIDATION OF FOUR SOCIAL SCALES FOR THE UX EVALUATION OF INTERACTIVE PRODUCTS

Ehsan Mortazavi, Philippe Doyon-Poulin, Daniel Imbeau, Jean-Marc Robert.

This article was submitted to the International Journal of Human-Computer Interaction on August 14, 2022

Abstract

The social dimension of interactive products covers all aspects of our relationships with others that are impacted by owning and using such products. Although social features are making their way into a growing number of interactive products, there is a lack of evaluation tool to capture the social dimension on the user experience (UX). In this study, we developed four social scales using the UEQ+ framework. For scale development, 229 participants rated their UX with products having social aspects. Exploratory factor analysis allowed us to identify four sub-dimensions (Identification, Social interaction, Social stimulation, and Social acceptance), each evaluated with four items. For scale validation, 450 participants evaluated the UX of three product categories, using the new social scales, AttrakDiff, and six UX dimensions of UEQ+. Results of MANOVA showed that the social scales discriminated the three categories ($F(8, 560) = 20.68$, $P < 0.001$, Pillai's trace = 0.456). The four social scales developed in this study can be combined with other UX dimensions of the UEQ+ modular framework to provide a comprehensive overview of user interaction with products.

Keywords: User experience, UX, UX evaluation, Social scale, UEQ+

6.1 Introduction

User experience (UX) is a multidimensional construct that encompasses all aspects of user interaction with products, services, and systems[30] [45]. It includes pragmatic dimensions, that are associated with products usability and functionality, as well as hedonic dimensions that are associated with users' psychological well-being, like pleasure and stimulation, that together give a comprehensive picture of user-product interaction [46]. As technology advances, new environments and ways of interactions such as virtual reality are introduced, resulting in new UX dimensions that should be considered in UX design and evaluation [197]. Despite the fact that people use more and more products that connect them to others, the social aspects of UX have not been adequately investigated in UX studies [197, 204]. Yet, they have major impacts at a personal level for communication and entertainment purposes, as well as at a business level to improve technology adoption and productivity [205]. Communicating with others, expressing oneself through the possession or use of a product, collaborating with others on a task, and creating an experience with others (i.e., co-experience) are all manifestations of social aspects.

The impact of social aspects on UX has been studied from two perspectives. The first characterizes it as an element of the context of use. Arhippainen and Tahti [42] identified social and cultural factors as part of the context in user-product interaction that influence UX. They mentioned time pressure, pressure to succeed or fail, and explicit and implicit requirements as examples of social factors [42]. Lallemand and Koenig [189] took a step further and developed the UX context scale (UXCS), a tool that can be used in conjunction with UX evaluation tools to portray the contexts within which users interact with a product. Other people's presence, interactions with them, and how users feel in that environment are items measured in UXCS under social context [189]. However, it does not evaluate the social aspects of the UX with a product.

The second perspective considers social aspect as a UX dimension that should be measured on its own. In order to ensure a common understanding, the terms used in this paper are classified as follows: at the highest level, a "dimension" represents a significant factor explaining the UX of a

[30] For the sake of brevity, hereafter, we only use product(s) instead of product(s), service(s), and system(s).

product [3], “sub-dimensions” are the constituent elements of each dimension and “items” are the constituent elements of a sub-dimension. Identification is an example of a UX sub-dimension of the social dimension present in the AttrakDiff questionnaire [6]. Identification is the degree to which a user relates to a product and expresses oneself through its possession [46]. It is described as either self-focused or relationship-focused. The former is concerned with the users' self-perception as a result of possessing or using a product, whereas the latter reflects their social identity [206]. Identification has roots in the influence and popularity needs of individuals, i.e., the need to be liked, respected, and regarded as influential to others [207], which is in line with Jordan’s notion of socio-pleasure [201]. In this study, we adopted the second perspective according to which social aspects are part of the UX evaluation tools.

6.2 Literature review

Reviewing the UX evaluation tools shows that social dimension is evaluated under different terms. Table 6.1 shows subjective UX evaluation tools and the social sub-dimensions they cover. General evaluation tools such as AttrakDiff are applicable to a variety of products, while others, such as the game experience questionnaire (GEQ), are developed for a particular field of use. However, neither covers all sub-dimensions of the social dimension.

Table 6.1 UX evaluation tools with social sub-dimensions

UX evaluation tools	Type	Social sub-dimensions	Number of items
AttrakDiff	General	Identification	7
meCUE	General	Status	3
Chinese UX questionnaire	General	Conformity, identification	5
EGameFlow	Specific (education)	Social interaction	6
Faser LX	Specific (education)	Subjective norm, self-image, interaction	5
GEQ	Specific (games)	Empathy, negative feeling, behavioral involvement	17
GUESS	Specific (games)	Social connection	4
MEEGA+	Specific (games)	Social interaction	3
Smart TV UX	Specific (smart TV)	Social relatedness	4

In addition to AttrakDiff, meCUE [8] is another generally applicable UX evaluation tool that includes a social sub-dimension. It is analogous to AttrakDiff in that it focuses on self-expression

and how a product can communicate identity to others under the UX sub-dimension of “status”. A more comprehensive view on social dimension of UX is presented by Park et al [32]. Although they did not develop a UX evaluation tool they defined sociability with three sub-dimensions — social emotion, social value, and friendship — for the UX of mobile phones and services. Sociability was defined as the “degree to which a product satisfies the user’s desire of being sociable”. The sub-dimensions respectively evaluate product’s ability to enable feeling and sharing emotions socially; its ability to support user values such as social issues and problems; and its ability to enable making relationship with other people. In another study, Ryu and Kim [208] evaluated sociability with only one item in their questionnaire (without covering the three sub-dimensions) to find whether or not a medical information system satisfied the user social needs. One can notice the limited attention to different social aspects in general UX evaluation tools. AttrakDiff and meCUE only evaluate identification, and the User Experience Questionnaire (UEQ) [7] — the most commonly used subjective evaluation tool — does not cover any social aspect [197]. The same goes for UEQ+ [50], a modular framework developed based on UEQ, including 20 UX dimensions that can be used for the evaluation of different products, yet none relates to social aspect.

Considering the impact of culture on UX, the Chinese UX questionnaire was developed based on AttrakDiff [148]. Its authors proposed “conformity” as the new UX dimension that addressed the eastern culture of giving importance to other’s opinions. Conformity deals with the prevalence of a product in the market and how widely it is used by people. It is to some extent close to social acceptance, another sub-dimension that has been rarely used in UX evaluation [97]. Social acceptance is “how a user feels when interacting with a system in relation to the social situation, e.g., how uncomfortable or embarrassed they feel with respect to other people or one's own norm” [97]. The negative impact of lack of social acceptability affects not just the UX of a product, but also the user's self and social image, with the risk of stigmatization and misjudgment by others [209]. Making an intimate relationship with others when watching smart TVs [127], sharing information via a mobile app [210], collaborating with other students to co-write a poem [211], and working in a multi-user virtual reality space are all examples of where different social sub-dimensions of UX need to be addressed.

Social interaction in the form of communicating with others, collaborating with teammates, or expressing oneself through game characters has been widely used in the game industry. Multiplayer online games, gaming conventions, and game streaming services are all examples of social interaction media for players and spectators. As a result, UX evaluation tools developed specifically for games include social dimension, such as GEQ [65] and Game User Experience Satisfaction Scale (GUESS) [212]. GEQ features a separate questionnaire for the social presence that covers players' emphatic responses, negative feelings towards others, and behavioral involvement in games [65]. Another field where social dimension has gained attention is education. Interactions within and between learners and instructors, subjective norm, and self-image are social sub-dimensions evaluated in the FASER LX tool that evaluates learners experience in an e-learning system [128]. Other researchers used serious games for learning purposes and developed tools like EGameFlow [213] and Model for the Evaluation of Educational Games (MEEGA) [70] with social sub-dimensions. Reviewing the items of these tools highlights the case-specific nature of their development. Moreover, there is no suggestion of the possibility of a modular use of these tools except for GEQ social presence module. As a result, it is unclear whether these tools can be applied to other contexts.

Analyzing the UX evaluation tools showed that the social dimension of UX concerns three categories: product, user, and group (society). The product-related items investigate whether a product can provide social interactions. For example in EGameFlow, the item "the game supports social interaction between players" highlights the product's capability of making interaction. The user-related items evaluate what users can do with the social features of a product. For instance, Smart TV UX questionnaire uses the item "I can form an intimate relationship with others by using the smart TV" and GEQ uses "I empathized with the other(s)" as actions that can be done by users with social features of products. The group-related items investigate how users/owners of a product are perceived by other people and social norms. The item "I was not worried about other people's judgement" of the flow sub-dimension of UX in IVE questionnaire [141] and the item "Rarely used – widely used" of the Chinese UX questionnaire are examples of this category. Regardless of the naming in different tools, the social sub-dimensions provided in Table 6.1 includes at least one of the three categories of items. However, none of the UX tools cover all the social sub-dimensions.

Clearly, the UX community needs a new tool to evaluate the social dimension of UX when interacting with a product that is not bounded to a specific application field (e.g., games or education) and ensure a comprehensive coverage. A promising way of achieving this goal is to develop social scales and include them into the UEQ+ framework. Modularity of the UEQ+ enables it to be used in different contexts, allowing the evaluator the flexibility to tailor the tool and evaluate the set of UX dimensions most relevant to the study [50]. Researchers can construct new UX dimensions and add them to the UEQ+ framework. For instance, response behavior, response quality, and comprehensibility are the new UX dimensions developed for voice assistants [96]. Similarly, haptic and acoustic dimensions were developed for the UX evaluation of household devices [95]. In UEQ+, each dimension consists of four items measuring the dimension and a single item that determines the importance of that dimension for the UX evaluation of the product. The evaluation of an item is done on a 7-point Likert scale with two semantic differential anchor points. The consistency of UEQ+ simplifies the UX evaluation compared to using a combination of UX evaluation tools each with different rating scales and possible overlapping dimensions [50]. It also provides an overall UX rating for products that can be used for comparisons. All these advantages make UEQ+ a suitable candidate for hosting UX social dimension that can be used with other UX dimensions for different products.

In this study, we designed, developed, and validated social scales that can be used with other UX dimensions of the UEQ+ framework. To this end, we followed the dimension development process used in [50] for other UX dimensions of the UEQ+ framework, focusing on social dimension and present the validated questionnaire.

6.3 Study framework

This study followed three phases: identification, integration, and validation (Figure 6.1). The **identification** phase consisted in extracting social sub-dimensions and corresponding items from the literature and from the analysis of the UX evaluation tools, and in selecting a sample of items for each social sub-dimension by UX experts. These items were used in the **integration** phase to develop social scales by performing Exploratory Factor Analysis (EFA) on the responses of 229 participants who rated their UX with products having social characteristics. The **validation** phase included another survey study with 450 participants who rated their UX with specific products

from three different product categories with various levels of social dimension. The results were used to calculate the reliability and validity of the social scales. This study was approved by Polytechnique Montreal's Research Ethics Board (CER-2122-47-D) and participants read and signed an informed consent form before taking part in the study.

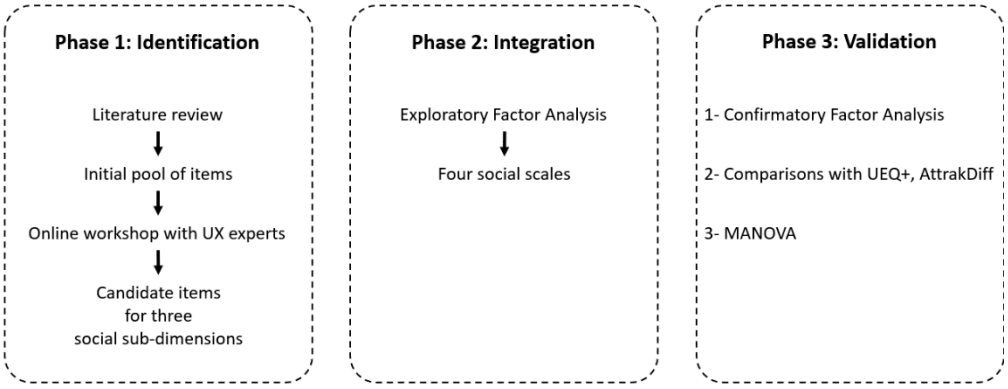

Figure 6.1 Three phases of the study

This paper is organized in eight sections. Section 4 presents the identification phase, followed by section 5 including the procedure, methodology, and results of the integration phase. Then, section 6 presents the validation phase and its main results with the validated social scales. Section 7 puts our results into perspective by comparing them with those of other social scales and presents limitations of our work. Section 8 concludes the paper with the main take-aways and some propositions for future research work.

6.3.1 Participants

Two types of participants took part in this study, UX experts in the identification phase and end users for the two rounds of data collection (integration and validation phases). Seven UX experts (two from the industry and five from academia) were brought together to work in an online workshop during the identification phase. They had from 5 to 25 years of work experience in the fields of human-computer interaction (HCI), UX, cognitive ergonomics, or digital accessibility.

End users who filled out questionnaires in the two rounds of data collection were from Canada and recruited through Amazon Mechanical Turk (AMT). The first survey included a sample of 229 participants who had experience in using an interactive product with social aspects. The second data collection captured users' experience with specific products from three product categories:

Social network, Online shopping, and Online banking. To that end, three separate questionnaires were prepared, and 150 participants responded to each category (total n=450). Table 6.2 presents the demographic data of experts and end user participants for each questionnaire in terms of age, gender, and level of education.

Table 6.2 Demographic data

Criteria		Online workshop	1st data collection	2nd data collection		
		UX experts (n=7)	Interactive products (n=229)	Social network (n=150)	Online shopping (n=150)	Online banking (n=150)
Age	18-29	0	70	23	32	33
	30-39	4	93	60	53	58
	40-49	1	38	38	31	29
	50-59	1	23	17	24	25
	60-69	1	5	11	10	4
	70+	0	0	1	0	1
Gender	Male	4	121	87	78	81
	Female	3	106	62	70	67
	Other	0	2	1	2	2
Education	High school or equivalent	0	41	39	35	43
	Bachelor	0	135	83	82	77
	MSc.	4	39	18	26	22
	Ph.D.	3	7	4	2	4
	Other	0	7	6	5	4

The 30-39 age group had the highest number of participants. For all questionnaires, participants were well balanced on gender. Overall, more than 70% of each group had a university-level education.

6.4 Phase 1 Identification

In a previous study, we performed a systematic literature review and an analysis of the UX evaluation tools and their dimensions [197]. As a result, we found that social dimension can be translated into three main sub-dimensions: identification, sociability, and social acceptance. **Identification** is the degree to which a user relates to a product and expresses oneself through its use or possession [46]. **Sociability** is the degree to which a product enables communication with others to meet user's social need [32]. **Social acceptance** is "how a user feels when interacting

with a system in relation to the social situation, e.g., how uncomfortable or embarrassed they feel with respect to other people or one's own norm" [97].

An online workshop with seven UX experts took place in December 2021 on the collaborative whiteboard platform Miro. They reviewed the three social sub-dimensions and 24 items, discussed the possibility of merging them, and proposed 12 new items in a 90-minute session. Results of the workshop were reassessed by four of these experts and the number of items was reduced from 36 to 27. Similar-meaning items were removed, and some modifications were made to the wording of the semantic differential poles. Overall, the 27 items were grouped under three social sub-dimensions: identification, sociability, and social acceptance.

6.5 Phase 2 Integration

6.5.1 Procedure

The first data collection with end users was done via an online questionnaire in a one-month period (March-April 2022). The questionnaire was designed on the QuestionPro website and data collection was done through AMT with participants receiving $2 compensation after having completed their response. The questionnaire was written in English and contained 34 questions. The first seven questions covered end users' demographic data, frequency of using products with social aspects, name of their selected product, and a brief description of the main usage of the product. The remaining 27 questions were on social items, among which there were seven items on identification, eight on social acceptance, and 12 on sociability. The 27 items on a 7-point semantic differential scale were presented in 5 sections, each with a short introductory sentence, similar to [50]. Participants were asked to evaluate their experience with products having social aspects. They were given examples of such products like social networks, messengers, forums, online games, and online collaboration tools. However, they were given the freedom to choose any social product with which they would frequently interact. Those who rarely or never used any of the selected products did not qualify to fill out the questionnaire. The average response time was 4 minutes and responses that took less than 2 minutes were removed. In addition to eliminating

incomplete answers, red-herring questions[31] were used in the questionnaires to catch participants who seemed to have randomly answered the questions. 148 participants (39%) were removed including 22 (6%) who did not meet the qualifications to answer the questionnaire, 57 (15%) who left the questionnaire incomplete, and 69 (18%) who failed the red-herring questions. Table 6.3 shows the variety of products with social aspects that were evaluated by participants.

Table 6.3 Frequency of the evaluated products of the first data collection

Product	Frequency (n=229)	Percentage (%)
Facebook	55	24.0
Instagram	25	10.9
Twitter	17	7.4
Reddit	15	6.6
Online games	12	5.2
Mobile devices	12	5.2
WhatsApp	7	3.1
YouTube	7	3.1
Discord	5	2.2
LinkedIn	5	2.2
Microsoft teams	5	2.2
Other	64	27.9

6.5.2 Methodology

Exploratory factor analysis (EFA) was used for refining and reducing the items of the first survey. Following the review by seven UX experts, we assumed that the candidate items describe only the sub-dimension they relate to. With a modular approach in mind for the social scales, we conducted 3 separates EFAs on identification (7 items), social acceptance (8 items), and sociability (12 items) instead of a single EFA on 27 items. However, the structural model is tested in the validation phase. As the initial step of EFA, we assessed the adequacy of sample size and the strength of correlations between items [199]. There are different suggestions for the suitable sample size for factor analysis. Some assert that the total number should be at least 300, while others suggest using the ratio of participants to items such as 10 to 1 or 5 to 1 [214]. We adopted the conservative approach of

[31] In questionnaires, red-herring or attention check questions are used to detect participants who do not read carefully or answer randomly.

having at least 200 respondents [215]. Suitability of data for factor analysis was explored by using Bartlett's test and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO). Multicollinearity was tested by checking whether the determinant of the correlation matrix of items is above 0.00001 and no pair of items has a correlation coefficient greater than 0.9. There should be a reasonable correlation between items measuring the same construct. Therefore, by investigating the correlation matrix, we removed the items displaying several correlations below 0.3 with other items [199].

We did factor extraction and rotation as the main analysis of EFA [199]. We applied principal axis factoring as the extraction method, and we used Kaiser's criterion, Scree plot, and parallel analysis to determine how many factors to retain. We used three criteria for factor extraction: factors with eigenvalues above Kaiser's criterion of 1, clear break on the Scree plot, and eigenvalues greater than the corresponding values for a randomly generated data matrix of the same size [214]. Factor analysis is an exploratory tool that assists researchers in making decisions based on different tests [199]. As a result, when criteria differ, decisions should be based on the knowledge of the researchers in respective fields. We applied the Varimax rotation technique and analyzed the rotated factor matrix for items with factor loadings above 0.3 [199]. Following the four items per UX dimension format of the UEQ+ framework [50], we kept the four items with highest factor loadings for each social sub-dimension. Lastly, the reliability of each scale was tested with Cronbach's alpha. Reliability refers to the internal consistency of the items of each factor. We used IBM SPSS 28 for statistical analyses.

6.5.3 Results

The result of the online workshop with experts was a sample of 27 items organized under five introductory sentences as shown in Table 6.4. They are at the basis of the questionnaire submitted to the end-user participants. The sub-dimension column shows the sub-dimension to which items used in each EFA belong. The sample size of 229 participants was adequate (i.e., n>200) for performing EFAs.

Table 6.4 List of items that were used in the first survey

No.	Items		Sub-dimensions
Q1	In my opinion, using this product in social situations makes me feel		
Q1-1	uncomfortable	comfortable	Social acceptance
Q1-2	embarrassed	proud	Social acceptance
Q1-3	rejected by others	accepted by others	Social acceptance
Q1-4	judged by others	approved by others	Social acceptance
Q2	In my opinion, using/owning this product makes me look		
Q2-1	cheap	stylish	Identification
Q2-2	out of date	modern	Identification
Q2-3	unprofessional	professional	Identification
Q2-4	old-fashioned	fashionable	Identification
Q2-5	odd	normal	Social acceptance
Q3	In my opinion, this product is		
Q3-1	unknown	widespread	Social acceptance
Q3-2	rarely used	widely used	Social acceptance
Q3-3	disliked by my peers	liked by my peers	Social acceptance
Q3-4	non-reflective of my values	reflective of my values	Identification
Q3-5	diminishing my standing (status)	enhancing my standing (status)	Identification
Q3-6	socially exclusive	socially inclusive	Sociability
Q3-7	dissuasive	persuasive	Sociability
Q3-8	socially neutral	socially empowering	Sociability
Q3-9	socially discouraging	socially encouraging	Sociability
Q3-10	socially disengaging	socially engaging	Sociability
Q3-11	non-valuable	valuable	Identification
Q4	In my opinion, using/owning this product		
Q4-1	does not support learning from interaction with others	supports learning from interaction with others	Sociability
Q4-2	rejects communities	supports communities	Sociability
Q5	In my opinion, by using this product I can		
Q5-1	be ignorant to others	be helpful to others	Sociability
Q5-2	not share my feelings	share my feelings	Sociability
Q5-3	isolate from others	communicate with others	Sociability
Q5-4	not collaborate with others	collaborate with others	Sociability
Q5-5	not empathize with others	empathize with others	Sociability

Investigating the correlation matrix of items of each EFA resulted in excluding two items (Q3-1 and Q3-2) for social acceptance and one item (Q3-7) for sociability. For these items, more than half of the correlation coefficients with other items were below 0.3. After their removal, the correlation matrices were recalculated. Results showed determinant values above the recommended value of 0.00001 for all the new correlation matrices. The KMO values exceeded the recommended value of 0.6 [214]. The Bartlett's test of Sphericity was statistically significant

for all sub-dimensions (at the alpha level of 0.01). Overall, all measures confirmed the suitability of the sample data and the absence of multicollinearity (Table 6.5).

Table 6.5 Measures for determining the suitability of sample data for EFAs

Measure		**Identification**	**Social acceptance**	**Sociability**
Determinant of correlation matrix		0.052	0.1	0.01
Kaiser-Meyer-Olkin measure of sampling adequacy		0.850	0.822	0.898
Bartlett's Test of Sphericity	Approx. Chi-Square	662.898	519.534	1025.175
	Degree of freedom	21	15	55
	Significance	<0.001	<0.001	<0.001

The Scree plot, parallel analysis, and Kaiser's criterion retained one factor for identification and one factor for social acceptance. It explained 55.2% of the variance for identification and 54.3% for social acceptance. However, there was not a consensus on the number of factors among the three criteria for sociability. Parallel analysis and scree plot showed one factor while Kaiser's criterion suggested retaining two. Considering the items that were loading on each factor, sociability was divided into two factors of social stimulation and social interaction. We defined social stimulation as qualities of a product that foster social interaction. Social interaction was also defined as the ways through which users make contact with others. The two-factor solution explained a total of 48.4 % of the variance. Social interaction accounted for 27.7% of the variance, whereas social stimulation contributed to 20.7 %.

Table 6.6 shows the rotated factor matrices of the 3 EFAs with corresponding factor loadings. The factor loadings represent the contribution each item makes to a factor [199]. The four items with the highest factor loading were selected for each social sub-dimension. These items had factor loading greater than 0.6, except for Q3-6. We tested content validity by investigating how theoretically relevant each item was to its factor which resulted in keeping Q3-6. This item was kept regardless of its relatively low loading on both factors, so as to conform with the 4-item format of the UEQ+ framework.

Table 6.6 Factor loading of items for the 3 EFAs

Identification		**Social acceptance**		**Sociability**		
Item number	Factor	Item number	Factor	Item number	Factor 1	Factor 2
Q2-4	**0.772**	Q1-4	**0.853**	Q5-3	**0.676**	
Q2-2	**0.750**	Q1-3	**0.827**	Q5-1	**0.649**	
Q2-1	**0.750**	Q1-2	**0.679**	Q5-5	**0.636**	
Q2-3	**0.681**	Q1-1	**0.646**	Q5-4	**0.618**	
Q3-5	0.663	Q2-5	0.567	Q4-2	0.607	
Q3-4	0.649	Q3-5	0.421	Q4-1	0.542	
Q3-11	0.559			Q5-2	0.533	
				Q3-9		**0.872**
				Q3-10	0.363	**0.652**
				Q3-8		**0.605**
				Q3-6	0.387	**0.388**

Cronbach' alpha exceeded the suggested cutoff of 0.7 confirming the reliability of all four social scales of identification, social acceptance, social interaction, and social stimulation as shown in Table 6.7.

Table 6.7 Reliability statistics of the four social scales

	Identification	**Social acceptance**	**Social interaction**	**Social stimulation**
Cronbach's alpha	0.841	0.846	0.802	0.792
Cronbach's alpha based on standardized items	0.844	0.845	0.803	0.793
Number of items	4	4	4	4

6.6 Phase 3 Validation

The goal of this phase was three-fold. First, to investigate the relationship between the social sub-dimensions using confirmatory factor analysis (CFA) and to test the fitness of this new data to the structural model we proposed. Second, to validate the scales by comparing the scores with those of other common evaluation tools. Third, to evaluate the social aspects of three product categories with the new social scales.

6.6.1 Procedure

The data collection was done over one month (May 2022) on the same survey and data collection websites and participants received $2 compensation after having completed their response. This time, three separate questionnaires, all written in English, were created with specific products to be evaluated from three different product categories. These products were Facebook, Instagram, and LinkedIn for the Social network product category; Amazon, Walmart, and eBay for the Online shopping product category; and CIBC, RBC, TD, and Scotiabank for the Online banking product category. We chose these three categories as they represented a range of socially engaging products, with social networks being highly social (e.g., messaging, community interactions), online shopping being moderately social (e.g., products posting, reviews, shared cart), and online banking being low on that continuum. Each questionnaire also contained the AttrakDiff questionnaire, six UX dimensions of the UEQ+ framework, and our newly developed social scales. This was done to investigate the validity of our social scales and the extent to which it is consistent with existing UX evaluation tools. Following the format of the UEQ+ framework, participants rated the importance of UX dimensions for the UX evaluation of each product category. Overall, participants answered a total of 89 questions. The average response time was 7 minutes for each questionnaire, and we removed responses that took less than 3 minutes. Similar data cleaning approach to the first survey was taken and 290 participants (i.e., 39%) were removed from which 49 (7%) participants did not meet the qualifications to answer the questionnaire, 90 (12%) left the questionnaire incomplete, and 151 (20%) failed red-herring questions resulting in a total of 450 responses. Table 6.8 shows the frequency of selection of the products evaluated.

Table 6.8 Frequency of selection of the products evaluated during the 2nd data collection (n=150 for each product category)

Product category	Name	Frequency	Percentage (%)
Social network	Facebook	98	65.3
	Instagram	44	29.3
	LinkedIn	8	5.4
Online shopping	Amazon	135	90
	Walmart	11	7.3
	eBay	4	2.7
Online banking	CIBC	25	16.7
	TD	30	20
	RBC	23	15.3
	Scotiabank	21	14
	Other	51	34

6.6.2 Methodology

Prior to conducting CFA for each product category, we used Mahanalobis distances to identify and remove multivariate outliers from the three samples collected in the second round of data collection [199]. Mahanalobis distances have a chi-square distribution and we used the recommended threshold value of $p<0.001$ for removing outliers [216]. As a result, 146, 142, and 144 valid responses were used for Social network, Online shopping, and Online banking product categories, respectively. We ran CFA for two models. The first model included four factors of identification, social interaction, social stimulation, and social acceptance. Considering that two factors of social interaction and social stimulation were two measures of the same underlying construct (i.e., sociability), large correlations between them were expected, which is why we used second-order CFA for sociability [217]. As a result, we did the CFA for the second model on three latent variables of identification, sociability, and social acceptance. We also used modification indices to improve the fitness of the model. CFA uses different goodness-of-fit indices to assess the quality of model fit to the data [215]. In this study, we reported: Chi-square statistical significance test, PCMIN/DF, Comparative Fit Index (CFI), Tucker Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). Composite Reliability (CR) was calculated, with values greater than 0.7 indicating acceptable reliability. Convergent validity that shows to what extent items of a factor are interrelated were tested by calculating Average Variance Extracted (AVE), with values above

0.5 deemed acceptable. The discriminant validity shows the extent to which latent factors are different and was tested by calculating AVE and Maximum Shared Variance (MSV) values [127].

To further investigate the validity of our scales, we analyzed the correlations of the four social scales with the scales of AttrakDiff and six scales of the UEQ+ framework to verify their redundancy with a threshold value of 0.8. As stated earlier, in addition to rating different items on a semantic differential scale, participants rated the importance of each dimension in the UX evaluation for the three product categories. We studied the ratings of social sub-dimensions to see whether they met our expectations of descending importance rating for Social network, Online shopping, and Online banking product categories.

The mean differences in ratings on the social scales between product categories were then examined using Multivariate Analysis of Variance (MANOVA). In this study the three product categories were considered as independent variables and the four social scales as the dependent variables. The data was checked to confirm that it met the MANOVA assumptions. The minimum sample size for conducting MANOVA is 20 responses per independent variable [198]. We tested the normality of the sample using the Shapiro-Wilk test. Multivariate normality was checked by calculating Mahanalobis distances [214]. We kept equal sample sizes across all independent variables [198]. We checked for the linearity between each pair of dependent variables that was tested using a scatterplot. Finally, multicollinearity was tested by measuring correlations among the dependent variables where values above 0.9 show potential for multicollinearity [214]. We used IBM SPSS 28 for statistical analyses and IBM SPSS Amos 28 for CFA.

6.6.3 Results

6.6.3.1 CFA

The results of the three CFAs of the first model showed a high correlation (0.92) between social interaction and social stimulation for the social network product category (Figure 6.2-left). According to EFA results, these two dimensions could be represented by a higher-level factor of sociability. Thus, using sociability as a second-order factor, we kept both social interaction and social stimulation in the second model (Figure 6.2-right). The three latent variables had correlation coefficients above 0.3 but below 0.9, indicating that they were reasonably related, but not

multicollinear. Although below the threshold, the correlation coefficient of 0.77 was observed between sociability and social acceptance.

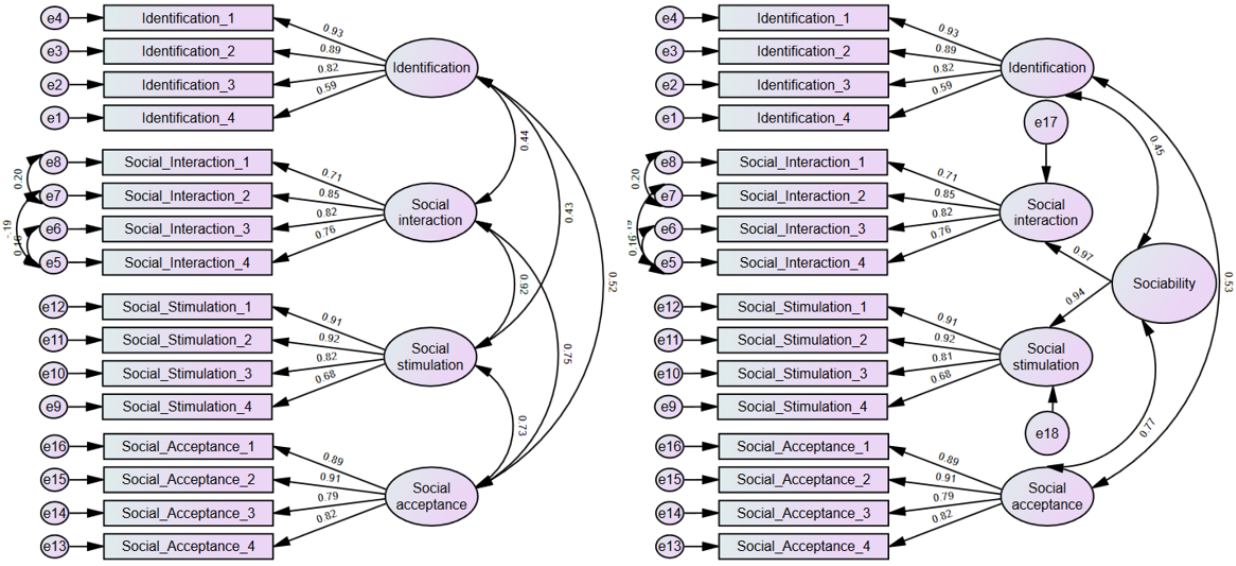

Figure 6.2 CFA for the first model with four factors (left) and the second model with three factors (right) for social network category

Table 6.9 Goodness of fit measures for the 3 CFAs of the second model shows the goodness of fit measures calculated for the three CFAs on the second model after the modifications. The thresholds shown in the table are from [217, 218]. Results confirm the model fit on the data for the three product categories.

Table 6.9 Goodness of fit measures for the 3 CFAs of the second model

Measure	**Threshold**	**Online banking**	**Online shopping**	**Social network**
Chi square	NA	154.385	162.808	208.459
Degree of freedom (df)	NA	95	93	96
Chi square/df	1-3	1.625	1.751	2.171
CFI	>0.95 (great), >0.90 (traditional)	0.959	0.956	0.939
TLI	>0.9	0.948	0.943	0.924
RMSEA	<0.5 (good) 0.5-.10 (moderate),>0.1 (bad)	0.066	0.073	0.090

Table 6.10 Factor loadings of the CFAs for the 3 product categories shows factor loadings greater than 0.5 for all items of each factor, which is better than the results of the EFAs (Table 6.6). The

average variance extract (AVE) of all factors exceeded the cutoff value of 0.5, confirming the convergent validity of the second model [198]. The discriminant validity of the three factors of identification, social acceptance, and sociability was confirmed because maximum shared variance (MSV) was greater than AVE for each factor, and the squared root of AVE (bold in Table 6.11) was greater than the correlation of the factors [219].

Table 6.10 Factor loadings of the CFAs for the 3 product categories

Factor	Item	Online banking	Online shopping	Social network
Identification	Identification_1	0.630	0.752	0.926
	Identification_2	0.670	0.849	0.890
	Identification_3	0.813	0.768	0.817
	Identification_4	0.784	0.844	0.588
Social interaction	Social_interaction_1	0.753	0.770	0.707
	Social_interaction_2	0.734	0.788	0.850
	Social_interaction_3	0.825	0.754	0.821
	Social_interaction_4	0.732	0.685	0.762
Social stimulation	Social_stimulation_1	0.841	0.910	0.911
	Social_stimulation_2	0.943	0.953	0.920
	Social_stimulation_3	0.653	0.665	0.815
	Social_stimulation_4	0.777	0.588	0.676
Social acceptance	Social_acceptance_1	0.929	0.951	0.891
	Social_acceptance_2	0.935	0.930	0.911
	Social_acceptance_3	0.704	0.716	0.787
	Social_acceptance_4	0.641	0.591	0.822
Sociability	Social interaction	0.993	0.932	0.972
	Social stimulation	0.776	0.872	0.942

Table 6.11 Model validity measures of the CFAs for the 3 product categories (AVE squared root in bold)

Category	Factor	Social acceptance	Identification	Sociability	CR	AVE	MSV
Online banking	Social acceptance	**0.813**			0.884	0.661	0.360
	Identification	0.600	**0.728**		0.817	0.530	0.360
	Sociability	0.427	0.347	**0.891**	0.884	0.794	0.182
Online shopping	Social acceptance	**0.815**			0.885	0.664	0.394
	Identification	0.562	**0.824**		0.894	0.679	0.316
	Sociability	0.628	0.499	**0.886**	0.880	0.785	0.394
Social network	Social acceptance	**0.854**			0.915	0.730	0.596
	Identification	0.525	**0.816**		0.886	0.666	0.276
	Sociability	0.772	0.454	**0.957**	0.956	0.916	0.596

Table 6.12 shows the final four social scales with the corresponding items used for the UX evaluation of the three product categories.

Table 6.12 Developed social scales and their items

Identification								
In my opinion, using/owning this product makes me look:								
Old-fashioned	1	2	3	4	5	6	7	Fashionable
Cheap	1	2	3	4	5	6	7	Stylish
Out of date	1	2	3	4	5	6	7	Modern
Unprofessional	1	2	3	4	5	6	7	Professional
Social interaction								
In my opinion, by using this product I can:								
Isolate from others	1	2	3	4	5	6	7	Communicate with others
Not collaborate with others	1	2	3	4	5	6	7	Collaborate with others
Be ignorant to others	1	2	3	4	5	6	7	Be helpful to others
Not empathize with others	1	2	3	4	5	6	7	Empathize with others
Social stimulation								
In my opinion, the use of this product is:								
Socially discouraging	1	2	3	4	5	6	7	Socially encouraging
Socially disengaging	1	2	3	4	5	6	7	Socially engaging
Socially neutral	1	2	3	4	5	6	7	Socially empowering
Socially exclusive	1	2	3	4	5	6	7	Socially inclusive
Social acceptance								
In my opinion, using this product in social situations makes me feel:								
Judged by others	1	2	3	4	5	6	7	Approved by others
Rejected by others	1	2	3	4	5	6	7	Accepted by others
Embarrassed	1	2	3	4	5	6	7	Proud
Uncomfortable	1	2	3	4	5	6	7	Comfortable

6.6.3.2 Comparisons

Table 6.13 shows the correlations between the four social scales developed in this study and the four scales of the AttrakDiff questionnaire for the three product categories. The highest correlations were observed with identification and attractiveness scales, particularly for the social network product category. All the statistically significant correlations are marked with a star in the table and those greater than 0.5 are highlighted in black.

Table 6.13 Correlations between the social scales and those of AttrakDiff (* significant with $P<0.05$)

Category	AttrakDiff scales	Identification	Social interaction	Social stimulation	Social acceptance

Social network	Pragmatic	0.331*	0.503*	0.549*	0.426*
	Identification	0.542*	0.648*	0.696*	0.584*
	Stimulation	0.534*	0.390*	0.427*	0.427*
	Attractiveness	0.622*	0.605*	0.665*	0.621*
Online shopping	Pragmatic	0.385*	0.326*	0.178*	0.433*
	Identification	0.502*	0.532*	0.456*	0.514*
	Stimulation	0.476*	0.263*	0.148	0.477*
	Attractiveness	0.521*	0.419*	0.295*	0.586*
Online banking	Pragmatic	0.326*	0.077	0.036	0.385*
	Identification	0.472*	0.347*	0.243*	0.479*
	Stimulation	0.412*	0.206*	0.137	0.401*
	Attractiveness	0.534*	0.188*	0.207*	0.507*

Table 6.14 shows the correlations between the four social scales developed in this study and the six scales of UEQ+. The four social scales were not redundant given that all the correlation coefficients were below 0.8. The highest correlations were observed with the Value and Attractiveness scales of the UEQ+. However, the social network product category displayed strong correlations with the pragmatic dimensions of Usefulness, Efficiency and Intuitive usage.

Table 6.14 Correlations between the social scales and the UEQ+ scales (* significant with P<0.05)

Category	UEQ+ scales	Identification	Social interaction	Social stimulation	Social acceptance
Social network	Efficiency	0.445*	0.508*	0.548*	0.457*
	Intuitive usage	0.296*	0.527*	0.555*	0.440*
	Perspicuity	0.276*	0.322*	0.393*	0.334*
	Value	0.694*	0.470*	0.522*	0.495*
	Usefulness	0.461*	0.696*	0.752*	0.619*
	Attractiveness	0.624*	0.658*	0.670*	0.645*
Online shopping	Efficiency	0.421*	0.296*	0.224*	0.404*
	Intuitive usage	0.347*	0.241*	0.169*	0.431*
	Perspicuity	0.339*	0.290*	0.095	0.361*
	Value	0.566*	0.471*	0.322*	0.542*
	Usefulness	0.476*	0.247*	0.206*	0.417*
	Attractiveness	0.544*	0.420*	0.306*	0.608*
Online banking	Efficiency	0.403*	0.043	0.028	0.429*
	Intuitive usage	0.331*	0.005	-0.020	0.456*
	Perspicuity	0.276*	-0.014	-0.033	0.407*
	Value	0.530*	0.206*	0.185*	0.434*
	Usefulness	0.384*	0.131	0.083	0.488*
	Attractiveness	0.459*	0.263*	0.213*	0.589*

Table 6.15 shows the average importance rating of each social scale. The highest ratings for all social scales went to social network and the lowest to online banking product category, as expected.

Table 6.15 Average importance rating of social scales for the three product categories (from 1: completely irrelevant to 7: very important)

Category	Identification	Social interaction	Social stimulation	Social acceptance
Social network	4.94	5.92	5.76	5.26
Online shopping	4.69	4.46	4.08	4.09
Online banking	4.02	3.22	3.06	3.27

6.6.3.3 MANOVA

Before conducting MANOVA to test the social scales ratings differences between product categories, participants who responded to more than one product category were removed yielding three independent samples of 106, 98, and 96 participants for social network, online shopping, and online banking product categories, respectively. Following Mahanalobis distance calculations, multivariate outliers were eliminated such that further analyses were performed on three independent samples of 95 participants.

Results of Shapiro-Wilk test of normality was significant ($P<0.001$), showing that the data was not normality distributed. However, the sample size was greater than 30 for each independent variable (i.e., product category), which is robust to the violation of normality or equality of variance [214]. The general pattern of scatterplot showed a linear relationship for each pair of dependent variables (i.e., four social scales). Table 6.16 shows that there was no multicollinearity between dependent variables because none of them had a correlation coefficient larger than 0.9 with another one. Moreover, being greater than 0.2, they were all at least moderately correlated. Box's test of equality of covariance matrices was significant ($P <0.001$) meaning that the observed covariance matrices of the dependent variables were not equal for all categories. Assumptions testing showed no violations for conducting MANOVA except for normality of data and homogeneity of variance-covariance matrices. As our sample size exceeds 30, violation of normality is not a concern. Moreover, the results of Pillai's trace was used for multivariate test because it is more robust when there is violation of assumptions [216].

Table 6.16 Correlation between the dependent variables (** correlation is significant at the 0.01 level)

Dependent variable	Identification	Social interaction	Social stimulation	Social acceptance
Identification	1	0.350**	0.399**	0.525**
Social interaction	0.350**	1	0.792**	0.485**
Social stimulation	0.399**	0.792**	1	0.480**
Social acceptance	0.525**	0.485**	0.480**	1

The results of one-way MANOVA showed significant difference between the three product categories on the combined social scales, F (8, 560) = 20.68, P< 0.001, Pillai's trace = 0.456, partial eta squared = 0.228. More investigation on each dependent variable showed statistical significance for social interaction and social stimulation using a Bonferroni adjusted alpha level of 0.0125 (highlighted in Table 6.17). Bonferroni adjustment is used to reduce the risk of a Type 1 error when several separate analyses are performed. Table 6.17 shows the multiple comparisons between product categories for each social scale.

Table 6.17 Significance of mean differences of dependent variables for three product categories (*significant at the 0.05 level and the highlighted are significant at 0.0125 level)

Dependent variable	Category (I)	Category(J)	Mean difference (I-J)	Std. Error	Sig.
Identification	Social network	Online shopping	0.684	0.1477	0.889
	Social network	Online banking	0.360*	0.1477	0.040
	Online shopping	Online banking	0.292	0.1477	0.120
Social interaction	Social network	Online shopping	1.455*	0.1532	<0.001
	Social network	Online banking	1.989*	0.1532	<0.001
	Online shopping	Online banking	0.534*	0.1532	0.002
Social stimulation	Social network	Online shopping	0.981*	0.1581	<0.001
	Social network	Online banking	1.639*	0.1581	<0.001
	Online shopping	Online banking	0.657*	0.1581	<0.001
Social acceptance	Social network	Online shopping	0.086	0.1460	0.823
	Social network	Online banking	0.402*	0.1460	0.017
	Online shopping	Online banking	0.315	0.1460	0.080

6.7 Discussion

This study covered the design, development, and validation of social scales for the UX evaluation of interactive products. UEQ+ scales used EFA to identify the four most representative items.

However, scale development in this study was different in that we also performed CFA on three different samples to provide more validity measures for the social scales. Results showed good internal consistency (Table 6.7), convergent validity, and discriminant validity (Table 6.10 and Table 6.11) for the social scales. Although we only have good model fit for the second model, the CFA confirmed our structural model of social dimension with the three main sub-dimensions of identification, sociability, and social acceptance.

Identification, social interaction, social stimulation, and social acceptance are the four social scales that were developed in this study (Table 6.12). The identification scale is similar to that of AttrakDiff and status dimension of meCUE. It deals with users' personal and social image. Users meet their influence and popularity need by expressing their identity, influence, and power to others through the use or possession of a product. The social interaction and social stimulation scales are complementary. The social stimulation scale represents the social incentives provided by a product to enable users to be socially active. A socially engaging, encouraging, empowering, and inclusive product enables users to achieve their social interaction goals in the form of communication, collaboration, and sharing of emotions with others. Finally, the social acceptance scale entails the feeling of being accepted by others. It can evoke self-conscious emotions like pride that are fulfilled when one is recognized and approved by others.

The social scales were subject to validation studies to find their relationship with other UX dimensions. Identification, the only social dimension of the AttrakDiff questionnaire, was expected to have strong correlations with the four social scales developed in this study. Our results confirmed this expectation by finding correlation coefficients between 0.45 and 0.69 with identification, except for social interaction ($r=0.34$) and social stimulation ($r=0.24$) for the online banking product category (Table 6.13). In addition to identification, we observed high correlations between the social scales and the attractiveness dimension of AttrakDiff. Hassenzahl [220] found a strong correlation between identification and beauty. He stated that beauty is social because it communicates identity and can be shared with and approved by others. Considering that beauty is an item in the attractiveness dimension of AttrakDiff, we can justify finding high correlations of social scales with attractiveness.

In addition to AttrakDiff, we measured the correlations of the four social scales with six UX dimensions of UEQ+. In the AttrakDiff, UEQ, and UEQ+ questionnaires, the attractiveness

dimension is used for the overall assessment of a product. Boos and Brau [95] calculated the correlations of the acoustic and haptic dimensions with the attractiveness dimension of UEQ+ to demonstrate that their new UX dimensions measured the same construct (i.e., UX). They concluded that the strong correlations with attractiveness validated the two new UX dimensions for the UEQ+ modular framework. Following the same approach, we found correlation coefficients ranging from 0.42 to 0.67 between the four social scales and the attractiveness dimension of UEQ+ (Table 6.14). However, there were three exceptions: social interaction (r=0.34) and social stimulation (r=0.24) for online banking, and social stimulation (r=0.30) for online shopping product category. We observed higher correlations between the social scales and the value dimension of UEQ+. This suggests that social scales are closer in nature to the hedonic dimensions of UX, such as value, than the pragmatic dimensions of efficiency or intuitive usage. However, high correlations with pragmatic dimensions were found in the social network product categories.

Validation of the social scales was tested by evaluating the UX of three product categories. These categories were selected to represent different levels of social aspects. The average importance rating of each scale in the UX evaluation of each product category followed our expectations, with social network having the highest importance ratings and online banking having the lowest across the four social scales (Table 6.15). Validation with three product categories addressed Lallemand and Koenig's [93] concerns on standardized scales development performing only a single validation study.

The MANOVA results showed a significant difference between the three product categories for the combination of social scales. Further analysis of each pair of categories revealed significant differences for social interaction and social stimulation (Table 6.17). The greatest difference was found in the social interaction scale between the social network and online banking product categories. It was to be expected, given that the former is designed as a communication platform, whereas the latter is exclusively task-related and rarely incorporates social interaction aspects. The mean differences for the two scales of identification and social acceptance were not significant for any product category when using a Bonferroni adjusted alpha level of 0.0125. The fact that all of the evaluated products were software applications could have had an impact on the results. It can be expected that the identification and social acceptance dimensions of software products do differ from those of physical products that attract more attention from users and spectators in a social

situation, such as virtual reality glasses. However, at the alpha level of 0.05, significant differences could be observed between social network and online banking, which are at the opposite extremes regarding the social features.

Overall, the four social scales developed in this study can be used in conjunction with other dimensions of the UEQ+ modular framework.

First limitation of this study is that the selected products were all software applications. Using a combination of software and hardware products could yield clear differences in social scales. Second limitation of this study is that we relied on the users' reported experience with the selected products. Using laboratory studies and doing specific tasks with better control over the hardware and context could provide more reliable results. However, this would have been possible only at a significantly increased cost with a likely much smaller sample.

6.8 Conclusion

In this study, we developed and validated four social scales for the UX evaluation of interactive products: identification, social interaction, social stimulation, and social acceptance. These scales are each measured with four items, following the same format as the other scales of the UEQ+ modular framework. They can be used in conjunction with the other dimensions of the framework for the UX evaluation of product with social aspects. Therefore, they have the benefits of UEQ+, such as being freely available, rapid to use, and providing a global score for easier comparisons. We used the results of an online workshop with UX experts to prepare the candidate items for each social scale. Then we did two rounds of data collection, the first used for scale development and the second for calculating reliability and the validity of scales. The scales were then validated through correlation analysis with the dimensions of the AttrakDiff questionnaire and the six UX dimensions of the UEQ+ framework. Results showed a high correlation with the identification and attractiveness dimensions of these two UX evaluation tools. Finally, using the social scales developed in this study we successfully discriminated among three product categories with different levels of social features. This study highlighted the importance of paying more attention to social dimensions. We recommend that future studies evaluate physical products with social dimensions to better validate the social scales.

CHAPTER 7 GENERAL DISCUSSION

The final objective of this thesis was to develop social scales within the context of UX evaluation through the fulfilment of three interrelated articles. We did not have "if you build it, they will come" approach for focusing on social dimension, rather it was driven by an actual need from the analysis of the literature. People are getting more dependent on different social products to present themselves in business context (e.g., LinkedIn) and personal context (e.g., Instagram, Tinder); to interact with others (e.g., Microsoft Teams, Skype); or to simply enjoy an experience like watching a movie (e.g., GroupWatch on Disney+). We began this research by performing a systematic literature review in the first article to update the state of the art on UX subjective evaluation tools. This decision was made because previous work used a limited number of databases [5, 82, 102], the studied time frame ended more than 5 years ago [5, 110], or focused on a subset of available questionnaires [9]. We identified 104 UX subjective evaluation tools in the form questionnaires and scales which is more comprehensive than 91 tools of [111] and 38 tools of [83]. We also categorized 300 UX dimensions into 13 categories and analyzed the identified UX tools accordingly. The findings showed that informational, social, physical, and cognitive dimensions were less frequently present in the current UX evaluation tools compared to usability, hedonic and utility dimensions. This finding is in line with what [54] and [44] found regarding the lower presence of the social and physical dimensions in users' stories of their UX with interactive products.

Two reasons led us to focus on the social dimension for the following articles. First, our results showed that mobile applications are the most widely evaluated products in UX studies, being used in different fields such as entertainment, learning and education, health, and work and training. Given that recent products with social features have been developed primarily on mobile platforms, addressing the social dimension can have an impact on a large spectrum of products. Second, we aimed not to be bound to any specific field or product category so that our findings can be generally applicable. Physical dimension is limited to physical products, cognitive dimension is limited to learning and training fields, informational dimension is limited to software products whereas social

dimension is present in both physical and software products in different fields of use. Clearly, social dimension seemed the most promising UX dimension to focus our work on.

The analysis of UX evaluation tools done as part of the SLR showed an emerging characteristic among the generally applicable tools developed in recent years: the adoption of a modular approach as done in meCUE and UEQ+. Due to their modularity, these tools are more flexible and better suited for product evaluation compared to tools with a fixed number of UX dimensions, some of which might not be relevant for the evaluation of certain products. In UX studies, the terms "relevance" and "importance" have been used interchangeably, but there is a subtle difference between them. The importance of UX dimensions ranges from extremely unimportant to extremely important. However when a dimension is meaningless for the evaluation of a product it is considered to be irrelevant. So, a UX dimension that is important is certainly relevant, but relevance does not imply importance. For instance, efficiency is very important and relevant for word processing products, whereas it is relevant but unimportant for games. Similarly, unimportance does not imply irrelevance. For instance, aesthetics can be unimportant for word processing products yet relevant to their UX. As discussed earlier assigning importance to UX dimensions is beneficial for designers and evaluators. It helps designers prioritize the most important dimensions when developing the new version of a product. It also assists novice evaluators in selecting the most important UX dimensions when using modular evaluation tools. Among UX evaluation tools, UEQ+ is the only tool that includes multiple UX dimensions as well as the option to rate importance (weight) of each UX dimension. Moreover, it is extendable, and researchers can construct new UX dimensions and add them to the UEQ+ framework. Analyzing the UEQ+ framework showed that the social dimension is not covered in any of the available scales. Thus, we decided to develop social scales to be added into the UEQ+ framework.

As people from different cultures have diverse perceptions, cognitions, and interaction styles with products [148], it is valid to question whether their UX evaluations differ regarding the importance ratings they assign to the UX dimension of different product categories. Santoso and Schrepp investigated the impact of culture on the importance ratings of UX dimensions for different software product categories in Germany and Indonesia [99]. The results of our SLR showed that 83% of the reviewed publications come from Europe and Asia. Therefore, for the second article, we decided to replicate the study of Santoso and Schrepp and extend it by adding five new UX

dimensions, including social dimension, ratings from end-users and UX experts, and having participants from Canada.

The results of the second article confirmed the results of [99] that the importance of UX dimensions for evaluating different categories of products was similar across different cultures. However, we observed that culture can affect the rating values for different UX dimensions. Our findings showed that Indonesians gave higher ratings for all UX dimensions, especially hedonic dimensions, compared to the other two cultures. We speculated that cultural differences influenced these ratings, and they can be explained through Hofstede's cultural dimensions. It is documented that countries with higher value on power distance dimensions of Hofstede are more cautious in freely expressing their opinions and giving negative feedback [192] and those with higher values on Hofstede's theory feminine dimension emphasize on the hedonic values [202]. Our results showed that UX experts considered hedonic and pragmatic dimensions as important while end users were more concerned with pragmatic dimensions. This finding is in line with those of [98], stating that experts take more UX aspects into consideration when asked to evaluate a product.

Regarding the social dimension, the results of the multi-dimensional scaling showed that identity, sociability, and social acceptance could be categorized into one group. This finding was important to us in that it provided insight into the potential sub-dimensions of the social dimension. The high importance of sociability, social acceptance, and identity for some product categories showed that the social dimension is required to be a part of the UX evaluation tools, particularly for product categories such as social networks. We also showed that the importance of UX dimensions can differ within the products of each category, calling for more attention when selecting UX dimensions for the evaluation of a product.

Considering the results of the first two articles, the final objective of this thesis was to develop social scales and add them into the UEQ+ framework. This was the subject of the third article. We developed four scales for the evaluation of the social dimension, being identification, social interaction, social stimulation, and social acceptance. These sub-dimensions cover different social aspects present in a user-product interaction. The identification scale is similar to that of AttrakDiff and the status dimension of meCUE. It deals with users' personal and social image. Users satisfy their need for influence and popularity by displaying their identity, influence, and power to others through the use or possession of a product. The social interaction and social stimulation scales are

complementary. The social stimulation scale represents the social incentives provided by a product to enable users to be socially active. A socially engaging, encouraging, empowering, and inclusive product enables users to achieve their social interaction goals in the form of communication, collaboration, and sharing of emotions with others. Finally, the social acceptance scale entails the feeling of being approved and accepted by others.

These scales are not bound to a specific product category, and one or more social sub-dimensions can be applicable depending on the evaluated product. They have the benefit of being rapid and easy to use, taking less than three minutes to complete and being used in conjunction with other UEQ+ scales. Lastly, validation of the social scales was tested by evaluating the UX of three product categories with varying levels of social aspects. Compared to the other studies that only developed new scales for UEQ+ [95, 96] without further validation, we have validated the scales with the responses from 450 users.

As pointed out by Väätäjä and Roto, UX methods that take time to be performed, such as longitudinal studies and large-scale experiments, are rarely used by industries [77]. Given this limitation, using evaluation tools with many items or a combination of tools with different formats and rating scales that may overlap on some dimensions is neither lean nor efficient. We showed in the SLR that questionnaires are still the most popular way to capture users' feedback, but they need to be updated to meet the fast-paced development cycle of new products. Modular evaluation tools and a granular approach towards evaluation of UX dimensions can solve these issues. With modular tools, format and rating scales are the same for all UX dimensions, and new scales would only be developed when new dimensions are defined. A granular approach to UX evaluation is to make sure that a product meets the most important UX dimensions first. As the ranking results of article 2 showed, pragmatic dimensions precede hedonic dimensions. Therefore, functionality and usability should be the priority for designers. The other approach is using the benchmarks provided by some UX evaluation tools (e.g., UEQ). However, further evaluations of different product categories with tools that measure the importance rating of each UX dimension can better determine the most important UX dimensions.

7.1 Limitations

The work in this thesis is subjected to two main limitations in addition to those presented in the three articles. The first limitation of this research is regarding the data collection. Common to any research using online data collection, participants might not be from the selected population, provide biased or random responses, and poorly represent the population by being more educated, having more access to the internet, or being more tech-savvy compared to an average person. The second limitation of this thesis is that, despite covering an important UX dimension, it contributes to the long list of UX dimensions by introducing the social dimension and its four sub-dimensions.

CHAPTER 8 CONCLUSION AND RECOMMENDATION

This research took place in the context of UX evaluation, which includes a large number of UX evaluation tools and UX dimensions. In order to have a clear understanding of the situation and address a significant gap in the literature, we defined the first objective of this thesis: providing the state of the art on UX evaluation tools and UX dimensions. We reached this objective in the first article by reviewing 325 studies from a list of 3831 papers. The categorization of 300 UX dimensions into 13 categories and the analysis of 104 UX evaluation tools led to the formulation of the next objectives. We selected the social dimension as the UX dimension that requires more attention and selected UEQ+ as the best tool for which we can develop social scales.

The downside of the modular evaluation tools such as UEQ+ is that the evaluators are responsible for the selection of UX dimensions. The second objective of this thesis was to address this challenge and investigate the importance of UX dimensions for different product categories. We fulfilled this objective in the second article by providing a list of important UX dimensions for different product categories. It was the result of a survey from 200 end users and 8 UX experts who evaluated the importance of 21 UX dimensions for 15 software product categories.

Developing and validating social scales was the third objective that was met in the third article. We analyzed the current tools and organized an online workshop with UX experts and came up with 27 items. Four social scales were identified after performing factor analysis. These scales each included four items with a single item to measure its importance. 450 respondents used these scales for the evaluation of three different product categories with varying level of social features. The analysis conformed to our expectations regarding the differences of social dimension between the three product categories.

Throughout this thesis we identified the social dimension as a less researched dimension, demonstrated its importance to different product categories, and developed and validated scales for its evaluation.

8.1 Contributions

Our research made four main contributions to the advancement of knowledge on UX evaluation. First, we classified the most recent evaluations tools based on the UX dimensions they measure and found that the social dimension lacked consideration from the UX community. This is of importance because social features are making their way into a growing number of interactive products. It is estimated that the individuals using social media will reach 4.4. billion worldwide by 2025 compared to 3.6 billion in 2020 [221].

Second, we confirmed that the importance rating of UX dimensions is dependent on the product categories rather than the culture, with new results from the Canadian population compared to the German and Indonesian populations. Moreover, we provided a list of important UX dimensions for different product categories. This is of importance because evaluators from different countries can use the same list of important UX dimensions when using modular evaluation tools. Similarly, designers can put more emphasis on the most important dimensions while meeting the culturally-specific design elements like language, currency, and date formatting.

Third, we developed and validated four scales to evaluate the social dimension of UX and incorporated them to the UEQ+ framework. It is important because it makes UEQ+ the only generally applicable tool that includes all aspects of the social dimension. The new social scales will benefit from the flexibility of the UEQ+ and can be used with other UX dimension for the evaluation of different products with social features.

We successfully applied and replicated the method developed by [50] to create an evaluation scale for new UX dimensions. Compared to previous scales of the UEQ+ [50, 95, 96], we used confirmatory factor analysis to show the relationship between the sub-dimensions we identified in the previous steps of scale development. Our SLR revealed that other UX dimensions are still neglected in their evaluation tools. Having a validated method to develop a new scale as part of the UEQ+ framework will be of great assistance to speed up the work of UX researchers.

8.2 Application to design

The mere presence of social scales can help designers by making them aware of different aspects of social dimension that could be considered in their design. Knowing that social interaction is

made through communication, collaboration, helping others, and empathizing with them can inspire designers to find a way to incorporate it in their design. As putting a “download pdf” icon beside a “print” icon seems to be unimportant but the eco-conscious attitude behind that button can change user behavior, similar thinking can encourage the introduction of new social features.

Ratings of social scales can be an indication to take action. For instance, high ratings of social interaction and stimulation with low rating of social acceptance can be alarming for social networks. According to the leaked internal documents from Facebook, there was an increase in users while more teenagers were experiencing mental health issues and left with low self-esteem and feeling of inadequacy. However, these findings were neglected from the company because in their opinion revenue trumped user mental health [222].

8.3 Future research works

We suggest that researchers and practitioners apply the new social scales for evaluating different software and hardware products that have social features. From smart jewelries to VR gaming glasses we expect to see different ratings for each social sub-dimension. We expect to see social scales be used partially or in whole in conjunction with other UX scales of the UEQ+ framework in future studies to provide feedback on how useful they can be.

Considering the benefits of the modular tools (flexible, consistent, easy to apply, fast, and extendable), future studies can use modular evaluation tools like UEQ+ framework and develop new scales to capture the underrepresented or new UX dimensions. Considering the current scales of the framework and the 13 categories provided in the first article, new studies can develop scales for engagement, physical characteristics, cognitive, and emotion dimensions.

REFERENCES

[1] E. L.-C. Law, V. Roto, M. Hassenzahl, A. P. Vermeeren, and J. Kort, "Understanding, scoping and defining user experience: a survey approach," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2009, pp. 719-728.

[2] J. McCarthy and P. Wright, "Technology as experience," *interactions,* vol. 11, no. 5, pp. 42–43, 2004, doi: 10.1145/1015530.1015549.

[3] J.-M. Robert, "Defining and structuring the dimensions of user experience with interactive products," in *11th International Conference on Engineering Psychology and Cognitive Ergonomics, EPCE 2014, Held as Part of 16th International Conference on Human-Computer Interaction, HCI International 2014, June 22, 2014 - June 27, 2014*, Heraklion, Crete, Greece, 2014, vol. 8532 LNAI: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 272-283, doi: 10.1007/978-3-319-07515-0_28. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-07515-0_28

[4] *ISO 9241-2019 (Ergonomics of human-system interaction)*, I. O. f. Standardization, 2019.

[5] J. Robinson, C. Lanius, and R. Weber, "The past, present, and future of UX empirical research," *Communication Design Quarterly Review,* vol. 5, no. 3, pp. 10-23, 2017.

[6] M. Hassenzahl, M. Burmester, and F. Koller, "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität," in *Mensch & Computer 2003: Interaktion in Bewegung*, G. Szwillus and J. Ziegler Eds. Wiesbaden: Vieweg+Teubner Verlag, 2003, pp. 187-196.

[7] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in *Symposium of the Austrian HCI and usability engineering group*, 2008: Springer, pp. 63-76.

[8] M. Minge, M. Thuring, I. Wagner, and C. V. Kuhr, "The meCUE questionnaire: A modular tool for measuring user experience," in *International Conference on Ergonomics Modeling, Usability and Special Populations, AHFE 2016, July 27, 2016 - July 31, 2016*, Walt Disney World, FL, United states, 2017, vol. 486: Springer Verlag, in Advances in Intelligent Systems and Computing, pp. 115-128, doi: 10.1007/978-3-319-41685-4_11. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-41685-4_11

[9] I. Díaz-Oreiro, G. López, L. Quesada, and L. A. Guerrero, "Standardized Questionnaires for User Experience Evaluation: A Systematic Literature Review," *Proceedings,* vol. 31, no. 1, p. 14, 2019. [Online]. Available: https://www.mdpi.com/2504-3900/31/1/14.

[10] D. Norman, J. Miller, and A. Henderson, "What you see, some of what's in the future, and how we go about doing it: HI at Apple Computer," in *Conference companion on Human factors in computing systems*, 1995, p. 155.

[11] M. Kurosu, "A proposal of the experience engineering (XE)," in *Joint 11th Asia Pacific Conference on Computer Human Interaction, APCHI 2013 and the 5th Indian Conference on Human Computer Interaction, India HCI 2013, September 24, 2013 - September 27, 2013*, Bangalore, India, 2013: Association for Computing Machinery, in ACM

International Conference Proceeding Series, pp. 108-113, doi: 10.1145/2525194.2525199. [Online]. Available: http://dx.doi.org/10.1145/2525194.2525199

[12] S. Rajeshkumar, R. Omar, and M. Mahmud, "Taxonomies of User Experience (UX) evaluation methods," in *2013 International Conference on Research and Innovation in Information Systems (ICRIIS)*, 27-28 Nov. 2013 2013, pp. 533-538, doi: 10.1109/ICRIIS.2013.6716765.

[13] Y. P. Cruz, C. A. Collazos, and T. Granollers, "The thin red line between usability and user experiences," in *Proceedings of the XVI International Conference on Human Computer Interaction*, 2015, pp. 1-2.

[14] C. Lallemand, G. Gronier, and V. Koenig, "User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey," *Computers in Human Behavior,* vol. 43, pp. 35-48, 2015.

[15] E. L.-C. Law, "The measurability and predictability of user experience," in *Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems*, 2011, pp. 1-10.

[16] J. A. Bargas-Avila and K. Hornbaek, "Old wine in new bottles or novel challenges? A critical analysis of empirical studies of User Experience," in *CHI 2011 - 29th Annual CHI Conference on Human Factors in Computing Systems, Conference Proceedings and Extended Abstracts*, 2011: Association for Computing Machinery, in Conference on Human Factors in Computing Systems - Proceedings, pp. 2689-2698, doi: 10.1145/1978942.1979336. [Online]. Available: http://dx.doi.org/10.1145/1978942.1979336

[17] J. Robinson and C. Lanius, "A Geographic and Disciplinary Examination of UX Empirical Research Since 2000," presented at the Proceedings of the 36th ACM International Conference on the Design of Communication, Milwaukee, WI, USA, 2018. [Online]. Available: https://doi.org/10.1145/3233756.3233930.

[18] M. Hassenzahl, A. Platz, M. Burmester, and K. Lehner, "Hedonic and ergonomic quality aspects determine a software's appeal," presented at the Proceedings of the SIGCHI conference on Human Factors in Computing Systems, The Hague, The Netherlands, 2000. [Online]. Available: https://doi.org/10.1145/332040.332432.

[19] S. Diefenbach, N. Kolb, and M. Hassenzahl, "The'hedonic'in human-computer interaction: history, contributions, and future research directions," in *Proceedings of the 2014 conference on Designing interactive systems*, 2014, pp. 305-314.

[20] D. L. Scapin, B. Senach, B. Trousse, and M. Pallot, "User experience: Buzzword or new paradigm?," in *5th International Conference on Advances in Computer-Human Interactions, ACHI 2012, January 30, 2012 - February 4, 2012*, Valencia, Spain, 2012: International Academy, Research and Industry Association, IARIA, in ACHI 2012 - 5th International Conference on Advances in Computer-Human Interactions, pp. 336-341.

[21] J. Park *et al.*, "Development of a web-based user experience evaluation system for home appliances," *International Journal of Industrial Ergonomics,* vol. 67, pp. 216-228, 2018, doi: 10.1016/j.ergon.2018.05.017.

[22] S. Hellweger and X. Wang, "What is user experience really: towards a UX conceptual framework," *arXiv preprint arXiv:1503.01850,* 2015.

[23] F. Lachner, P. Naegelein, R. Kowalski, M. Spann, and A. Butz, "Quantified UX: Towards a common organizational understanding of user experience," in *9th Nordic Conference on Human-Computer Interaction, NordiCHI 2016, October 23, 2016 - October 27, 2016*, Gothenburg, Sweden, 2016, vol. 23-27-October-2016: Association for Computing Machinery, in ACM International Conference Proceeding Series, pp. Chalmers University of Technology; et al.; IT-University of Gothenburg; Tobii Pro; University of Gothenburg; Visage Technologies, doi: 10.1145/2971485.2971501. [Online]. Available: http://dx.doi.org/10.1145/2971485.2971501

[24] A. Allam and H. M. Dahlan, "User experience: Challenges and opportunities," *Journal of Information Systems Research and Innovation,* vol. 3, no. 1, pp. 28-36, 2013.

[25] J. Redish and C. Barnum, "Overlap, Influence, Intertwining: The Interplay of UX and Technical Communication," *J. Usability Studies,* vol. 6, no. 3, pp. 90–101, 2011.

[26] M. Zarour and M. Alharbi, "User experience framework that combines aspects, dimensions, and measurement methods," *Cogent Engineering,* vol. 4, no. 1, Dec 2017, Art no. 1421006, doi: 10.1080/23311916.2017.1421006.

[27] A. Følstad and R. K. Rolfsen, "Measuring the effect of User Experience design changes in e-Commerce web sites: A case on customer guidance," in *User Experience–Towards a UniÞed View: Proceedings of the 2nd COST294-MAUSE International Open Workshop*, 2006, pp. 10-15.

[28] N. Bevan, "Extending Quality in Use to Provide a Framework for Usability Measurement," in *Human Centered Design*, Berlin, Heidelberg, M. Kurosu, Ed., 2009// 2009: Springer Berlin Heidelberg, pp. 13-22.

[29] L. Gegner, M. Runonen, and T. Keinonen, "Oscillating between extremes: a framework for mapping differing views on User eXperience," presented at the Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces, Milano, Italy, 2011. [Online]. Available: https://doi.org/10.1145/2347504.2347566.

[30] N. Fenton and S. L. Pflegger, *Software metrics: a rigorous and practical approach*, Second edition ed. PWS Publishing Company, 1997.

[31] V. Roto, "User experience from product creation perspective," *Towards a UX Manifesto,* vol. 31, p. 34, 2007.

[32] J. Park, S. H. Han, H. K. Kim, Y. Cho, and W. Park, "Developing elements of user experience for mobile phones and services: Survey, interview, and observation approaches," *Human Factors and Ergonomics In Manufacturing,* vol. 23, no. 4, pp. 279-293, 2013, doi: 10.1002/hfm.20316.

[33] V. Roto, E.-C. Law, A. P. Vermeeren, and J. Hoonhout, "User experience white paper: Bringing clarity to the concept of user experience," 2011.

[34] J. Kalbach, *Mapping experiences*, 2nd ed. O'Reilly Media, 2020.

[35] J.-M. Robert and A. Lesage, "From usability to user experience with interactive systems," in *The Handbook of Human-Machine Interaction. A Human-centered Design Approach*, 2011, pp. 303-320.

[36] L. Alben, "Quality of experience: defining the criteria for effective interaction design," *interactions,* vol. 3, no. 3, pp. 11-15, 1996.

[37] M. Hassenzahl and N. Tractinsky, "User experience-a research agenda," *Behaviour & information technology,* vol. 25, no. 2, pp. 91-97, 2006.

[38] N. Ariza and J. Maya, "Towards an empirical model of the UX: A factor analysis study," in *9th International Conference on Design and Emotion 2014, October 6, 2014 - October 10, 2014*, Bogota, Cali and Medellin, Colombia, 2014: Universidad de Los Andes, in 9th International Conference on Design and Emotion 2014: The Colors of Care, pp. 689-697.

[39] M. Hassenzahl, "User experience (UX): towards an experiential perspective on product quality," presented at the Proceedings of the 20th Conference on l'Interaction Homme-Machine, Metz, France, 2008. [Online]. Available: https://doi.org/10.1145/1512714.1512717.

[40] K. Schulze and H. Krömker, "A framework to measure user experience of interactive online products," presented at the Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research, Eindhoven, The Netherlands, 2010. [Online]. Available: https://doi.org/10.1145/1931344.1931358.

[41] D. Norman and J. Nielsen. https://www.nngroup.com/articles/definition-user-experience/ (accessed August 2022.

[42] L. Arhippainen and M. Tähti, "Empirical evaluation of user experience in two adaptive mobile application prototypes," in *Mum 2003. proceedings of the 2nd international conference on mobile and ubiquitous multimedia*, 2003, no. 011: Citeseer, pp. 27-34.

[43] E. Karapanos, J. Zimmerman, J. Forlizzi, and J.-B. Martens, "User experience over time: an initial framework," presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 2009. [Online]. Available: https://doi.org/10.1145/1518701.1518814.

[44] G. Provost and J.-M. Robert, "The dimensions of positive and negative user experiences with interactive products," in *International Conference of Design, User Experience, and Usability*, 2013: Springer, pp. 399-408.

[45] J.-M. Robert and A. Lesage, "Designing and evaluating user experience," in *The Handbook of Human-Machine Interaction: A Human-Centered Design Approach*, G. A. Boy Ed. UK: Ashgate, 2011, ch. 15, pp. 321-338.

[46] M. Hassenzahl, "The Thing and I: Understanding the Relationship Between User and Product," in *Funology: From Usability to Enjoyment*, M. A. Blythe, K. Overbeeke, A. F. Monk, and P. C. Wright Eds. Dordrecht: Springer Netherlands, 2003, pp. 31-42.

[47] P. Desmet and P. Hekkert, "Framework of product experience," *International journal of design,* vol. 1, no. 1, 2007.

[48] M. Thüring and S. Mahlke, "Usability, aesthetics and emotions in human–technology interaction," *International Journal of Psychology,* vol. 42, no. 4, pp. 253-264, 2007/08/01 2007, doi: 10.1080/00207590701396674.

[49] V. Roto and M. Rautava, "User experience elements and brand promise," in *International engagability & design conference, in conjunction with NordiCHI*, 2008, vol. 8: Citeseer.

[50] M. Schrepp and J. Thomaschewski, "Design and Validation of a Framework for the Creation of User Experience Questionnaires," *International Journal of Interactive Multimedia & Artificial Intelligence,* vol. 5, no. 7, 2019.

[51] P. G. Zimmermann, "Beyond usability: Measuring aspects of user experience," PhD, ETH Zurich, 2008.

[52] P. Wright, J. McCarthy, and L. Meekison, "Making sense of experience," in *Funology*: Springer, 2003, pp. 43-53.

[53] J. Kort, A. Vermeeren, and J. E. Fokker, "Conceptualizing and measuring user experience," *Towards a UX manifesto,* p. 57, 2007.

[54] J.-M. Robert and A. Larouche, "The dimensions of user experience with interactive systems," in *IADIS International Conference Interfaces and Human Computer Interaction 2012 (part of MCCSIS 2012)*, 2012, pp. 89-96.

[55] M. Rochefeuille, "Analyse des dimensions de l'expérience utilisateur avec des produits interactifs," Département de mathématiques et de génie industriel, École Polytecvhnique de Montréal, 2013. [Online]. Available: https://publications.polymtl.ca/1159/

[56] T. Orehovački, "Evaluating the Interplay of User Experience Facets in the Context of Social Web Applications," presented at the Proceedings of the 2020 9th International Conference on Software and Information Engineering (ICSIE), Cairo, Egypt, 2020. [Online]. Available: https://doi.org/10.1145/3436829.3436863.

[57] A. Larouche, "Enquête visant à déterminer les dimensions de l'expérience utilisateur," Master's projet report, Département de mathématiques et de génie industriel, École Polytechnique de Montréal, 2011.

[58] G. Provost, "Études des expériences des utilisateurs avec des produits interactifs," Master's projet report, Département de mathématiques et de génie industriel, École Polytechnique de Montréal, 2012.

[59] M. Konstantakis, K. Michalakis, J. Aliprantis, E. Kalatha, and G. Caridakis, "Formalising and evaluating Cultural User Experience," in *2017 12th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, 9-10 July 2017 2017, pp. 90-94, doi: 10.1109/SMAP.2017.8022675.

[60] M. Toomim, T. Kriplean, C. Pörtner, and J. Landay, "Utility of human-computer interactions: Toward a science of preference measurement," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011, pp. 2275-2284.

[61] M. Moshagen and M. T. Thielsch, "Facets of visual aesthetics," *International Journal of Human-Computer Studies,* vol. 68, no. 10, pp. 689-709, 2010/10/01/ 2010, doi: https://doi.org/10.1016/j.ijhcs.2010.05.006.

[62] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the PANAS scales," *Journal of personality and social psychology,* vol. 54, no. 6, p. 1063, 1988.

[63] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavior Therapy and Experimental Psychiatry,* vol. 25, no. 1, pp. 49-59, 1994/03/01/ 1994, doi: https://doi.org/10.1016/0005-7916(94)90063-9.

[64] K. Poels, Y. A. de Kort, and W. A. IJsselsteijn, "D3. 3: Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games," 2007.

[65] W. A. IJsselsteijn, Y. A. de Kort, and K. Poels, "The game experience questionnaire," *Eindhoven: Technische Universiteit Eindhoven,* vol. 46, no. 1, 2013.

[66] C. Shu-Hui, W. Wann-Yih, and J. Dennison, "Validation of EGameFlow: A self-report scale for measuring user experience in video game play," *Computers in Entertainment,* vol. 16, no. 3, 2018, doi: 10.1145/3238249.

[67] A. M. Moosa, N. Al-Maadeed, M. Saleh, S. A. Al-Maadeed, and J. M. Aljaam, "Designing a Mobile Serious Game for Raising Awareness of Diabetic Children," *IEEE Access,* vol. 8, pp. 222876-222889, 2020, doi: 10.1109/ACCESS.2020.3043840.

[68] M. Skjuve and P. B. Brandzaeg, "Measuring user experience in chatbots: An approach to interpersonal communication competence," in *5th International Conference on Internet Science, INSCI 2018, October 24, 2018 - October 26, 2018*, St. Petersburg, Russia, 2019, vol. 11551 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 113-120, doi: 10.1007/978-3-030-17705-8_10. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-17705-8_10

[69] L. G. R. De Lima, A. De Lima Salgado, and A. P. Freire, "Evaluation of the user experience and Intrinsic motivation with educational and mainstream digital games," in *7th Latin American Conference on Human Computer Interaction, CLIHC 2015, November 18, 2015 - November 21, 2015*, Cordoba, Argentina, 2015: Association for Computing Machinery, in Proceedings of the 7th Latin American Conference on Human Computer Interaction, CLIHC 2015, doi: 10.1145/2824893.2824904. [Online]. Available: http://dx.doi.org/10.1145/2824893.2824904

[70] G. Petri, C. G. von Wangenheim, and A. F. Borgatto, "MEEGA+: an evolution of a model for the evaluation of educational games," *INCoD/GQS,* vol. 3, pp. 1-40, 2016.

[71] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Advances in psychology*, vol. 52: Elsevier, 1988, pp. 139-183.

[72] B. Keshavarz and H. Hecht, "Validating an efficient method to quantify motion sickness," *Human factors,* vol. 53, no. 4, pp. 415-426, 2011.

[73] P. Kourtesis, S. Collina, L. A. Doumas, and S. E. MacPherson, "Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology," *Frontiers in human neuroscience,* vol. 13, p. 417, 2019.

[74] C. M. MacDonald and M. E. Atwood, "Changing perspectives on evaluation in HCI: past, present, and future," presented at the CHI '13 Extended Abstracts on Human Factors in Computing Systems, Paris, France, 2013. [Online]. Available: https://doi.org/10.1145/2468356.2468714.

[75] N. H. Basri, N. L. M. Noor, W. A. W. Adnan, F. M. Saman, and A. H. A. Baharin, "Conceptualizing and understanding user experience," in *2016 4th International Conference on User Science and Engineering (i-USEr)*, 23-25 Aug. 2016 2016, pp. 81-84, doi: 10.1109/IUSER.2016.7857938.

[76] A. P. Vermeeren, E. L.-C. Law, V. Roto, M. Obrist, J. Hoonhout, and K. Väänänen-Vainio-Mattila, "User experience evaluation methods: current state and development needs," in *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries*, 2010, pp. 521-530.

[77] H. Väätäjä and V. Roto, "Questionnaires in user experience evaluation," in *UXEM workshop held in conjunction with INTERACT*, 2009, vol. 24, no. 28.8: Citeseer.

[78] M. Holman and F. Adebesin, "Taking the Subjectivity out of UX Evaluation with Emotiv EPOC+," presented at the Proceedings of the South African Institute of Computer Scientists and Information Technologists 2019, Skukuza, South Africa, 2019. [Online]. Available: https://doi.org/10.1145/3351108.3351139.

[79] S. Federici, M. L. Mele, M. Bracalenti, A. Buttafuoco, R. Lanzilotti, and G. Desolda, "Bio-behavioral and self-report user experience evaluation of a usability assessment platform (UTAssistant)," in *3rd International Conference on Human Computer Interaction Theory and Applications, HUCAPP 2019 - Part of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISIGRAPP 2019, February 25, 2019 - February 27, 2019*, Prague, Czech republic, 2019, vol. 2: SciTePress, in VISIGRAPP 2019 - Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, pp. 19-27.

[80] S. Soleimani and E. L.-C. Law, "What Can Self-Reports and Acoustic Data Analyses on Emotions Tell Us?," presented at the Proceedings of the 2017 Conference on Designing Interactive Systems, Edinburgh, United Kingdom, 2017. [Online]. Available: https://doi.org/10.1145/3064663.3064770.

[81] N. Ibrahim, W. Fatimah, W. Ahmad, and A. F. Shafie, "User Experience Study on Folktales Mobile Application for Children's Education," in *9th International Conference on Next Generation Mobile Applications, Services and Technologies, NGMAST 2015, September 9, 2015 - September 11, 2015*, Cambridge, United kingdom, 2015: Institute of Electrical and Electronics Engineers Inc., in Proceedings - NGMAST 2015: The 9th International Conference on Next Generation Mobile Applications, Services and Technologies, pp. 353-358, doi: 10.1109/NGMAST.2015.73. [Online]. Available: http://dx.doi.org/10.1109/NGMAST.2015.73

[82] C. L. B. Maia and E. S. Furtado, "A systematic review about user experience evaluation," in *5th International Conference on Design, User Experience, and Usability, DUXU 2016 and Held as Part of 18th International Conference on Human-Computer Interaction, HCI International 2016, July 17, 2016 - July 22, 2016*, Toronto, ON, Canada, 2016, vol. 9746: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes

in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 445-455, doi: 10.1007/978-3-319-40409-7_42. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-40409-7_42

[83] V. Roto *et al.* http://www.allaboutux.org/ (accessed 27 July 2022).

[84] L. Rivero and T. Conte, "A Systematic Mapping Study on Research Contributions on UX Evaluation Technologies," presented at the Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems, Joinville, Brazil, 2017. [Online]. Available: https://doi.org/10.1145/3160504.3160512.

[85] V. Montreuil, "Vers un prototype d'outil d'évaluation subjective de l'expérience utilisateur avec des produits interactifs," Département de mathématiques et de génie industriel, Polytechnique Montreal, 2014. [Online]. Available: https://publications.polymtl.ca/1381/

[86] M. Hassenzahl and A. Monk, "The inference of perceived usability from beauty," *Human–Computer Interaction,* vol. 25, no. 3, pp. 235-260, 2010.

[87] T. Walsh, P. Nurkka, J. Varsaluoma, H. Petrie, S. Kujala, and C. Power, "Axe UX: Exploring long-term user experience with iScale and AttrakDiff," in *18th International Academic MindTrek Conference, MINDTREK 2014, November 4, 2014 - November 6, 2014*, Tampere, Finland, 2014: Association for Computing Machinery, in MINDTREK 2014 - Proceedings of the 18th International Academic MindTrek Conference:, pp. 32-39, doi: 10.1145/2676467.2676480. [Online]. Available: http://dx.doi.org/10.1145/2676467.2676480

[88] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and evaluation of a short version of the user experience questionnaire (UEQ-S)," *International Journal of Interactive Multimedia and Artificial Intelligence, 4 (6), 103-108.,* 2017.

[89] A. Hinderks, M. Schrepp, F. J. Dominguez Mayo, M. J. Escalona, and J. Thomaschewski, "Developing a UX KPI based on the user experience questionnaire," *Computer Standards and Interfaces,* vol. 65, pp. 38-44, 2019, doi: 10.1016/j.csi.2019.01.007.

[90] M. Schrepp, J. r. Thomaschewski, and A. Hinderks, "Construction of a benchmark for the user experience questionnaire (UEQ)," *Int. J. Interact. Multimed. Artif. Intell.,* vol. 4, pp. 40-44, 2017.

[91] M. Minge, L. Riedel, and M. Thüring, "Modulare Evaluation interaktiver Technik. Entwicklung und Validierung des meCUE Fragebogens zur Messung der User Experience," *Grundlagen und Anwendungen der Mensch-Technik-Interaktion,* vol. 10, pp. 28-36, 2014.

[92] M. Minge, M. Thuring, and I. Wagner, "Developing and validating an English version of the meCUE questionnaire for measuring user experience," in *Human Factors and Ergonomics Society 2016 International Annual Meeting, HFES 2016, September 19, 2016 - September 23, 2016*, Washington, DC, United states, 2016, vol. 0: Human Factors and Ergonomics Society Inc., in Proceedings of the Human Factors and Ergonomics Society, pp. 2056-2060, doi: 10.1177/1541931213601468. [Online]. Available: http://dx.doi.org/10.1177/1541931213601468

[93] C. Lallemand and V. Koenig, ""How could an intranet be like a friend to me?" - Why standardized UX scales don't always fit," in *35th Annual Conference of the European Association of Cognitive Ergonomics, ECCE 2017, September 20, 2017 - September 22,*

2017, Umea, Sweden, 2017, vol. Part F131193: Association for Computing Machinery, in ACM International Conference Proceeding Series, pp. 9-16, doi: 10.1145/3121283.3121288. [Online]. Available: http://dx.doi.org/10.1145/3121283.3121288

[94] M. Minge and M. Thuring, "The MeCUE questionnaire (2.0): Meeting five basic requirements for lean and standardized UX assessment," in *7th International Conference on Design, User Experience, and Usability, DUXU 2018 Held as Part of HCI International 2018, July 15, 2018 - July 20, 2018*, Las Vegas, NV, United states, 2018, vol. 10918 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 451-469, doi: 10.1007/978-3-319-91797-9_33. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-91797-9_33

[95] B. Boos and H. Brau, "Erweiterung des UEQ um die Dimensionen Akustik und Haptik," in *Mensch und Computer 2017 - Usability Professionals*, S. Hess and H. Fischer Eds.: Gesellschaft für Informatik e.V., 2017.

[96] A. M. Klein, M. Schrepp, A. Hinderks, and J. Thomaschewski, "Measuring User Experience Quality of Voice Assistants Voice Communication Scales for the UEQ plus Framework," in *2020 15th Iberian Conference on Information Systems and Technologies*, A. Rocha, B. E. Perez, F. G. Penalvo, M. D. Miras, and R. Goncalves Eds., (Iberian Conference on Information Systems and Technologies, 2020.

[97] B. Weiss and S. Willkomm, "User experience and social attribution for an embodied spoken dialog system," in *Workshop on Computers As Social Actors, CASA 2013 - Co-located with 13th International Conference on Intelligent Virtual Agents, IVA 2013, September 28, 2013*, Edinburgh, United kingdom, 2013, vol. 1119: CEUR-WS, in CEUR Workshop Proceedings, pp. 1-8.

[98] A. Jaaskelainen and K. Heikkinen, "Divergence of user experience: Professionals vs. end users," in *1st International Workshop on the Interplay Between User Experience and Software Development, I-UxSED 2010, October 17, 2010 - October 17, 2010*, Reikjavick, Ireland, 2010, vol. 656: CEUR-WS, in CEUR Workshop Proceedings, pp. 8-12.

[99] H. B. Santoso and M. Schrepp, "The impact of culture and product on the subjective importance of user experience aspects," *Heliyon,* vol. 5, no. 9, Sep 2019, Art no. e02434, doi: 10.1016/j.heliyon.2019.e02434.

[100] J. Mashapa and D. Van Greunen, "User experience evaluation metrics for usable accounting tools," in *2010 Annual Research Conference of the South African Institute of Computer Scientist and Information Technologists, SAICSIT 2010, October 11, 2010 - October 13, 2010*, Bela Bela, South africa, 2010: Association for Computing Machinery, in ACM International Conference Proceeding Series, pp. 170-181, doi: 10.1145/1899503.1899522. [Online]. Available: http://dx.doi.org/10.1145/1899503.1899522

[101] V. Roto, M. Obrist, and K. Väänänen-Vainio-Mattila, "User experience evaluation methods in academic and industrial contexts," in *Proceedings of the Workshop UXEM*, 2009, vol. 9: Citeseer, pp. 1-5.

[102] I. Pettersson, F. Lachner, A.-K. Frison, A. Riener, and A. Butz, "A Bermuda Triangle? A Review of Method Application and Triangulation in User Experience Evaluation," presented at the Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal QC, Canada, 2018. [Online]. Available: https://doi.org/10.1145/3173574.3174035.

[103] K. Vaananen-Vainio-Mattila, T. Olsson, and J. Hakkila, "Towards Deeper Understanding of User Experience with Ubiquitous Computing Systems: Systematic Literature Review and Design Framework," in *Human-Computer Interaction - Interact 2015, Pt Iii*, vol. 9298, J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, and M. Winckler Eds., (Lecture Notes in Computer Science, 2015, pp. 384-401.

[104] A. C. Ten and F. Paz, "A systematic review of user experience evaluation methods in information driven websites," in *6th International Conference on Design, User Experience, and Usability, DUXU 2017, held as part of the 19th International Conference on Human-Computer Interaction, HCI 2017, July 9, 2017 - July 14, 2017*, Vancouver, BC, Canada, 2017, vol. 10288 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 492-506, doi: 10.1007/978-3-319-58634-2_36. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-58634-2_36

[105] W. T. Nakamura, E. H. T. De Oliveira, and T. Conte, "Usability and user experience evaluation of learning management systems a systematic mapping study," in *19th International Conference on Enterprise Information Systems, ICEIS 2017, April 26, 2017 - April 29, 2017*, Porto, Portugal, 2017, vol. 3: SciTePress, in ICEIS 2017 - Proceedings of the 19th International Conference on Enterprise Information Systems, pp. 97-108, doi: 10.5220/0006363100970108. [Online]. Available: http://dx.doi.org/10.5220/0006363100970108

[106] A. T. Revi, D. E. Millard, and S. E. Middleton, "A Systematic Analysis of User Experience Dimensions for Interactive Digital Narratives," in *13th International Conference on Interactive Digital Storytelling, ICIDS 2020, November 3, 2020 - November 6, 2020*, Bournemouth, United kingdom, 2020, vol. 12497 LNCS: Springer Science and Business Media Deutschland GmbH, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 58-74, doi: 10.1007/978-3-030-62516-0_5. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-62516-0_5

[107] G. C. Guerino and N. M. C. Valentim, "Usability and user experience evaluation of natural user interfaces: a systematic mapping study," *Iet Software,* vol. 14, no. 5, pp. 451-467, Oct 2020, doi: 10.1049/iet-sen.2020.0051.

[108] Y. M. Kim, I. Rhiu, and M. H. Yun, "A Systematic Review of a Virtual Reality System from the Perspective of User Experience," *International Journal of Human-Computer Interaction,* vol. 36, no. 10, pp. 893-910, Jun 2020, doi: 10.1080/10447318.2019.1699746.

[109] L. Rivero and T. Conte, "A systematic mapping study on research contributions on UX evaluation technologies," in *16th Brazilian Symposium on Human Factors in Computing Systems, IHC 2017, October 23, 2017 - October 27, 2017*, Joinville, SC, Brazil, 2017: Association for Computing Machinery, in ACM International Conference Proceeding Series, p. Brazilian Computer Society; et al.; Federal University of Parana (UFPR); Federal

University of Sao Carlos (UFSCar); Santa Catarina State University (UDESC); The Open University, doi: 10.1145/3160504.3160512. [Online]. Available: http://dx.doi.org/10.1145/3160504.3160512

[110] M. Zarour and M. Alharbi, "user experience aspects and dimensions: Systematic literature review," *International Journal of Knowledge Engineering,* vol. 3, no. 2, pp. 52-59, 2017.

[111] T. Darin, B. Coelho, and B. Borges, "Which Instrument Should I Use? Supporting Decision-Making About the Evaluation of User Experience," in *8th International Conference on Design, User Experience, and Usability, DUXU 2019, held as part of the 21st International Conference on Human-Computer Interaction, HCI International 2019, July 26, 2019 - July 31, 2019*, Orlando, FL, United states, 2019, vol. 11586 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 49-67, doi: 10.1007/978-3-030-23535-2_4. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-23535-2_4

[112] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.

[113] K. Yue, D. L. Wang, X. P. Yang, H. C. Hu, Y. Q. Liu, and X. Q. Zhu, "Evaluation of the user experience of "astronaut training system": An Immersive, VR-Based, Motion-Training System," in *Optical Measurement Technology and Instrumentation*, vol. 10155, S. Han and J. Tan Eds., (Proceedings of SPIE, 2016.

[114] C. Rossato, P. Pluchino, N. Cellini, G. Jacucci, A. Spagnolli, and L. Gamberini, "Facing with Collaborative Robots: The Subjective Experience in Senior and Younger Workers," *Cyberpsychology Behavior and Social Networking,* 2021, doi: 10.1089/cyber.2020.0180.

[115] J. Schneider, D. Borner, P. Van Rosmalen, and M. Specht, "Presentation trainer, your public speaking multimodal coach," in *ACM International Conference on Multimodal Interaction, ICMI 2015, November 9, 2015 - November 13, 2015*, Seattle, WA, United states, 2015: Association for Computing Machinery, in ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction, pp. 539-546, doi: 10.1145/2818346.2830603. [Online]. Available: http://dx.doi.org/10.1145/2818346.2830603

[116] G. Cajamarca, I. Rodriguez, V. Herskovic, and M. Campos, "StraightenUp: Implementation and Evaluation of a Spine Posture Wearable," in *Ubiquitous Computing and Ambient Intelligence, Ucami 2017*, vol. 10586, S. F. Ochoa, P. Singh, and J. Bravo Eds., (Lecture Notes in Computer Science, 2017, pp. 655-665.

[117] R. Lorenz, J. Pascual, B. Blankertz, and C. Vidaurre, "Towards a holistic assessment of the user experience with hybrid BCIs," *Journal of Neural Engineering,* vol. 11, no. 3, 2014, doi: 10.1088/1741-2560/11/3/035007.

[118] H.-Y. Chang, W.-J. Tseng, C.-E. Tsai, H.-Y. Chen, R. L. Peiris, and L. Chan, "Facepush: Introducing normal force on face with head-mounted displays," in *31st Annual ACM Symposium on User Interface Software and Technology, UIST 2018, October 14, 2018 - October 17, 2018*, Berlin, Germany, 2018: Association for Computing Machinery, in UIST 2018 - Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, pp. 927-935, doi: 10.1145/3242587.3242588. [Online]. Available: http://dx.doi.org/10.1145/3242587.3242588

[119] A. Min, D. Lee, G. Gao, S. Jeong, and P. C. Shih, "Design and assessment of a personal breathalyzer intervention to support responsible drinking," *International Journal of Human-Computer Studies,* vol. 137, May 2020, Art no. 102382, doi: 10.1016/j.ijhcs.2019.102382.

[120] R. Hammady, M. H. Ma, and A. Powell, "User Experience of Markerless Augmented Reality Applications in Cultural Heritage Museums: 'MuseumEye' as a Case Study," in *Augmented Reality, Virtual Reality, and Computer Graphics, Avr 2018, Pt Ii*, vol. 10851, L. T. DePaolis and P. Bourdot Eds., (Lecture Notes in Computer Science, 2018, pp. 349-369.

[121] J. Cuadros, V. Serrano, J. Garcia-Zubia, and U. Hernandez-Jayo, "Design and evaluation of a user experience questionnaire for remote labs," *IEEE Access,* pp. 1-1, 2021, doi: 10.1109/ACCESS.2021.3069559.

[122] T. J. Saun and T. P. Grantcharov, "Development of the User Experience (UX) and Video Quality Evaluation (VQE) Instruments for Assessment of Intraoperative Video Capture Technology," *Journal of Surgical Education,* vol. 78, no. 1, pp. 201-206, Jan-Feb 2021, doi: 10.1016/j.jsurg.2020.06.011.

[123] B. Barbot and J. C. Kaufman, "What makes immersive virtual reality the ultimate empathy machine? Discerning the underlying mechanisms of change," *Computers in Human Behavior,* vol. 111, Oct 2020, Art no. 106431, doi: 10.1016/j.chb.2020.106431.

[124] C. Boletsis, "A user experience questionnaire for vr locomotion: Formulation and preliminary evaluation," in *7th International Conference on Augmented Reality, Virtual Reality, and Computer Graphics, AVR 2020, September 7, 2020 - September 10, 2020*, Lecce, Italy, 2020, vol. 12242 LNCS: Springer Science and Business Media Deutschland GmbH, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 157-167, doi: 10.1007/978-3-030-58465-8_11. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-58465-8_11

[125] V. Nagalingam, R. Ibrahim, and R. C. M. Yusoff, "EDUGXQ: User Experience Instrument for Educational Games' Evaluation," *International Journal of Advanced Computer Science and Applications,* vol. 11, no. 1, pp. 562-569, Jan 2020. [Online]. Available: <Go to ISI>://WOS:000518467600070.

[126] N. Pretto, E. Micheloni, S. Gasparotto, C. Fantozzi, G. De Poli, and S. Canazza, "Technology-Enhanced Interaction with Cultural Heritage: An Antique Pan Flute from Egypt," *Acm Journal on Computing and Cultural Heritage,* vol. 13, no. 2, Jun 2020, Art no. 8, doi: 10.1145/3355395.

[127] J. Jang and M. Y. Yi, "Determining and validating smart TV UX factors: A multiple-study approach," *International Journal of Human Computer Studies,* vol. 130, pp. 58-72, 2019, doi: 10.1016/j.ijhcs.2019.05.001.

[128] Y. Safsouf, K. Mansouri, and F. Poirier, "Design of a new scale to measure the learner experience in e-learning systems," in *International Conference on e-Learning 2019, EL 2019, July 17, 2019 - July 19, 2019*, Porto, Portugal, 2019: IADIS Press, in Multi Conference on Computer Science and Information Systems, MCCSIS 2019 - Proceedings of the International Conference on e-Learning 2019, pp. 301-304, doi:

10.33965/el2019_201909c042. [Online]. Available: http://dx.doi.org/10.33965/el2019_201909c042

[129] E. Ormeño, C. Alessio, S. F. Ochoa, and C. A. Collazos, "Towards a formative instrument to evaluate user experience in virtual reality serious games," presented at the Proceedings of the 5th Workshop on ICTs for improving Patients Rehabilitation Research Techniques, Popayan, Columbia, 2019. [Online]. Available: https://doi.org/10.1145/3364138.3364152.

[130] S. Karunakaran and E. Brorson, "Spam in User Generated Content Platforms: Developing the HaBuT Instrument to Measure User Experience," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 6-9 Oct. 2019 2019, pp. 1981-1988, doi: 10.1109/SMC.2019.8914165.

[131] J. Moizer *et al.*, "An approach to evaluating the user experience of serious games," *Computers and Education,* vol. 136, pp. 141-151, 2019, doi: 10.1016/j.compedu.2019.04.006.

[132] I. Rodriguez, M. Karyda, A. Lucero, and V. Herskovic, "Aestimo: A Tangible Kit to Evaluate Older Adults' User Experience," in *Human-Computer Interaction - Interact 2019, Pt I*, vol. 11746, D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, and P. Zaphiris Eds., (Lecture Notes in Computer Science, 2019, pp. 13-32.

[133] E. V. Wilson and S. Djamasbi, "Measuring mobile user experience instruments for research and practice," *Communications of the Association for Information Systems,* vol. 44, no. 1, 2019, doi: 10.17705/1CAIS.04408.

[134] M. Yu, R. Zhou, H. Wang, and W. Zhao, "An evaluation for VR glasses system user experience: The influence factors of interactive operation and motion sickness," *Applied Ergonomics,* vol. 74, pp. 206-213, 2019, doi: 10.1016/j.apergo.2018.08.012.

[135] C. Alberola, G. Walter, and H. Brau, "Creation of a short version of the user experience questionnaire UEQ," 2018, vol. 17: Walter de Gruyter GmbH, in i-com, 1 ed., pp. 57-64, doi: 10.1515/icom-2017-0032. [Online]. Available: http://dx.doi.org/10.1515/icom-2017-0032

[136] S. Nielsen *et al.*, "Subjective Experience of Interacting with a Social Robot at a Danish Airport," in *2018 27th Ieee International Symposium on Robot and Human Interactive Communication*, J. J. Cabibihan, F. Mastrogiovanni, A. K. Pandey, S. Rossi, and M. Staffa Eds., (Ieee Ro-Man, 2018, pp. 1163-1170.

[137] T. Orehovaki, D. Plantak Vukovac, M. Deko, and Z. Stapi, "Evaluating relevant UX dimensions with respect to iot ecosystem intended for students activities tracking and success prediction," in *5th International Conference on Learning and Collaboration Technologies, LCT 2018 Held as Part of HCI International 2018, July 15, 2018 - July 20, 2018*, Las Vegas, NV, United states, 2018, vol. 10924 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 279-293, doi: 10.1007/978-3-319-91743-6_22. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-91743-6_22

[138] S.-C. Yang, T.-L. Lee, and T.-T. Feng, "User Experience of Mobile Application's Interface: Measurement Development," presented at the Proceedings of the 5th Multidisciplinary

International Social Networks Conference, Saint-Etienne, France, 2018. [Online]. Available: https://doi.org/10.1145/3227696.3227698.

[139] P. Krawczyk, M. Topolewski, and M. Pallot, "Towards a reliable and valid mixed methods instrument in user eXperience studies," in *23rd International Conference on Engineering, Technology and Innovation, ICE/ITMC 2017, June 27, 2017 - June 29, 2017*, Madeira Island, Portugal, 2017, vol. 2018-January: Institute of Electrical and Electronics Engineers Inc., in 2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017 - Proceedings, pp. 1455-1464, doi: 10.1109/ICE.2017.8280054. [Online]. Available: http://dx.doi.org/10.1109/ICE.2017.8280054

[140] S. T. Lee and M. Ivory, "What makes a good UX questionnaire? User-centered scorecard development," in *International Conference on Affective and Pleasurable Design, 2016, July 27, 2016 - July 31, 2016*, Walt Disney World, FL, United states, 2017, vol. 483: Springer Verlag, in Advances in Intelligent Systems and Computing, pp. 281-293, doi: 10.1007/978-3-319-41661-8_28. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-41661-8_28

[141] K. Tcha-Tokey, O. Christmann, E. Loup-Escande, and S. Richir, "Proposition and Validation of a Questionnaire to Measure the User Experience in Immersive Virtual Environments," *International Journal of Virtual Reality,* vol. 16, no. 1, pp. 33-48, 01/01 2016, doi: 10.20870/IJVR.2016.16.1.2880.

[142] C. Roth and H. Koenitz, "Evaluating the user experience of interactive digital narrative," in *1st ACM International Workshop on Multimedia Alternate Realities, AltMM 2016, October 16, 2016*, Amsterdam, Netherlands, 2016: Association for Computing Machinery, Inc, in AltMM 2016 - Proceedings of the 1st International Workshop on Multimedia Alternate Realities, co-located with ACM Multimedia 2016, pp. 31-36, doi: 10.1145/2983298.2983302. [Online]. Available: http://dx.doi.org/10.1145/2983298.2983302

[143] J. S. Laine and T. Leppnen, "Experimental comparison of the user experiences of different digital and printed newspaper versions," *Journal of Print and Media Technology Research,* vol. 5, no. 2, pp. 107-131, Jun-Jul 2016, doi: 10.14622/jpmtr-1603.

[144] E. Cavalcante, L. Rivero, and T. Conte, "MAX: A method for evaluating the post-use user eXperience through cards and a board," in *27th International Conference on Software Engineering and Knowledge Engineering, SEKE 2015, July 6, 2015 - July 8, 2015*, Pittsburgh, PA, United states, 2015, vol. 2015-January: Knowledge Systems Institute Graduate School, in Proceedings of the International Conference on Software Engineering and Knowledge Engineering, SEKE, pp. 495-500, doi: 10.18293/SEKE2015-136. [Online]. Available: http://dx.doi.org/10.18293/SEKE2015-136

[145] V. Macedo and C. Marcio Silva, "Building a semantic differential scale as tool for assisting UX evaluation with home appliances," in *3rd International Conference on Design, User Experience, and Usability: Theories, Methods, and Tools for Designing the User Experience, DUXU 2014, Held as Part of 16th International Conference on Human-Computer Interaction, HCI International 2014, June 22, 2014 - June 27, 2014*, Heraklion, Crete, Greece, 2014, vol. 8517 LNCS: Springer Verlag, in Lecture Notes in Computer

Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), PART 1 ed., pp. 308-317, doi: 10.1007/978-3-319-07668-3_30. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-07668-3_30

[146] G. Regal, P. Wokerstorfer, M. Busch, M. Tscheligi, and C. Hochleitner, "TACTUX - A tactile user experience assessment board," in *8th Nordic Conference on Human-Computer Interaction, NordiCHI 2014, October 26, 2014 - October 30, 2014*, Helsinki, Finland, 2014: Association for Computing Machinery, in Proceedings of the NordiCHI 2014: The 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, pp. 983-986, doi: 10.1145/2639189.2670256. [Online]. Available: http://dx.doi.org/10.1145/2639189.2670256

[147] R. Bernhaupt and M. Pirker, "Evaluating user experience for interactive television: Towards the development of a domain-specific user experience questionnaire," in *14th IFIP TC 13 International Conference on Human-Computer Interaction, INTERACT 2013, September 2, 2013 - September 6, 2013*, Cape Town, South africa, 2013, vol. 8118 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), PART 2 ed., pp. 642-659, doi: 10.1007/978-3-642-40480-1_45. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40480-1_45

[148] S. Liu, X. S. Zheng, G. Liu, J. Jian, and K. Peng, "Beautiful, usable, and popular: Good experience of interactive products for Chinese users," *Science China Information Sciences,* vol. 56, no. 5, pp. 1-14, 2013, doi: 10.1007/s11432-013-4835-4.

[149] A. Sonnleitner, M. Pawlowski, T. Kasser, and M. Peissner, "Experimentally manipulating positive user experience based on the fulfilment of user needs," in *14th IFIP TC 13 International Conference on Human-Computer Interaction, INTERACT 2013, September 2, 2013 - September 6, 2013*, Cape Town, South africa, 2013, vol. 8120 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), PART 4 ed., pp. 555-562, doi: 10.1007/978-3-642-40498-6_45. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40498-6_45

[150] M. Korber, A. Eichinger, K. Bengler, and C. Olaverri-Monreal, "User experience evaluation in an automotive context," in *2013 IEEE Intelligent Vehicles Symposium, IEEE IV 2013, June 23, 2013 - June 26, 2013*, Gold Coast, QLD, Australia, 2013: Institute of Electrical and Electronics Engineers Inc., in IEEE Intelligent Vehicles Symposium, Proceedings, pp. 13-18, doi: 10.1109/IVS.2013.6629440. [Online]. Available: http://dx.doi.org/10.1109/IVS.2013.6629440

[151] J. Zhuo, D. Huang, and Y. Chen, "Exploring user experience measurement of location-based service applications," in *19th International Conference on Industrial Engineering and Engineering Management: Management System Innovation, October 27, 2012 - October 29, 2012*, Changsha, China, 2013: Springer Science and Business Media Deutschland GmbH, in 19th International Conference on Industrial Engineering and Engineering Management: Management System Innovation, pp. 969-978, doi: 10.1007/978-3-642-38427-1-102. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-38427-1-102

[152] R. Buchner, D. Wurhofer, A. Weiss, and M. Tscheligi, "User experience of industrial robots over time," in *7th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI'12, March 5, 2012 - March 8, 2012*, Boston, MA, United states, 2012: Association for Computing Machinery, in HRI'12 - Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction, pp. 115-116, doi: 10.1145/2157689.2157712. [Online]. Available: http://dx.doi.org/10.1145/2157689.2157712

[153] S. Kujala, V. Roto, K. Vaananen-Vainio-Mattila, E. Karapanos, and A. Sinnela, "UX Curve: A method for evaluating long-term user experience," *Interacting with Computers,* vol. 23, no. 5, pp. 473-483, 2011, doi: 10.1016/j.intcom.2011.06.005.

[154] B. Van De Laar, H. Gurkok, D. Plass-Oude Bos, F. Nijboer, and A. Nijholt, "Perspectives on user experience evaluation of brain-computer interfaces," in *Universal Access in Human-Computer Interaction: Users Diversity - 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011, Proceedings*, 2011, vol. 6766 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), PART 2 ed., pp. 600-609, doi: 10.1007/978-3-642-21663-3_65. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-21663-3_65

[155] L. Wang and M. Lv, "Study on assessing user experience of augmented reality applications," in *12th International Conference on Virtual, Augmented and Mixed Reality, VAMR 2020, held as part of the 22nd International Conference on Human-Computer Interaction, HCII 2020, July 19, 2020 - July 24, 2020*, Copenhagen, Denmark, 2020, vol. 12190 LNCS: Springer, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 208-222, doi: 10.1007/978-3-030-49695-1_14. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-49695-1_14

[156] S. Wang, Y. Liu, F. Song, X. Xie, and D. Yu, "Research on Evaluation System of User Experience With Online Live Course Platform," *IEEE Access,* vol. 9, pp. 23863-23875, 2021, doi: 10.1109/ACCESS.2021.3054047.

[157] Z. H. Huang, H. D. Li, and Iop, "Research on UX Evaluation Model of Computer Input Devices (Keyboard and Mouse) Based on AHP," in *4th International Conference on Design, Mechanical and Material Engineering*, vol. 686, (IOP Conference Series-Materials Science and Engineering, 2019.

[158] J. Lin, S. J, and Y. C, "Research on User Experience Evaluation Model of Smart Jewelry Based on Kansei Engineering," in *8th International Conference on Design, User Experience, and Usability, DUXU 2019, held as part of the 21st International Conference on Human-Computer Interaction, HCI International 2019, July 26, 2019 - July 31, 2019*, Orlando, FL, United states, 2019, vol. 11586 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 164-171, doi: 10.1007/978-3-030-23535-2_12. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-23535-2_12

[159] J. Liu, M. Cui, and L. Jia, "Evaluation of user experience based 3D websites using gray correlation analysis and AHP," in *2018 14th International Conference on Natural*

Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 28-30 July 2018 2018, pp. 1306-1309, doi: 10.1109/FSKD.2018.8687167.

[160] C. J. Van Staden, J. A. Van Biljon, and J. H. Kroeze, "EModeration: Towards a User Experience Evaluation Framework," in *2015 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists, SAICSIT 2015, September 28, 2015 - September 30, 2015*, Stellenbosch, South africa, 2015, vol. 28-30-September-2015: Association for Computing Machinery, in ACM International Conference Proceeding Series, p. Amazon Development Centre Cape Town; IBM South Africa; Information Science Department and the Centre for AI Research at Stellenbosch University; Meraka Institute of the Council for Scientific and Industrial Research (CSIR); Microsoft South Africa; OutSystems South Africa, doi: 10.1145/2815782.2815821. [Online]. Available: http://dx.doi.org/10.1145/2815782.2815821

[161] J. P. Wan and Y. H. Zhu, "Research on smart mobile phone user experience with grounded theory," in *2013 International Conference on Information Systems and Computing Technology, ISCT 2013, September 15, 2013 - September 16, 2013*, Wuxi, China, 2013: CRC Press/Balkema, in Information Systems and Computing Technology - Proceedings of the International Conference on Information Systems and Computing Technology, ISCT 2013, pp. 51-58, doi: 10.1201/b15789-9. [Online]. Available: http://dx.doi.org/10.1201/b15789-9

[162] F. G. Faust *et al.*, "Mixed prototypes for the evaluation of usability and user experience: simulating an interactive electronic device," *Virtual Reality,* vol. 23, no. 2, pp. 197-211, Jun 2019, doi: 10.1007/s10055-018-0356-1.

[163] F. Freitas, H. Oliveira, I. Winkler, and M. Gomes, "Virtual Reality on Product Usability Testing: A Systematic Literature Review," in *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, 7-10 Nov. 2020 2020, pp. 67-73, doi: 10.1109/SVR51698.2020.00025.

[164] J. Vuorio, J. Okkonen, and J. Viteli, "Finnish upper secondary students user expectations and experiences using mall system," in *22nd International Academic Mindtrek Conference and the Mindtrek Conference, Mindtrek 2018, October 10, 2018 - October 11, 2018*, Tampere, Finland, 2018: Association for Computing Machinery, in ACM International Conference Proceeding Series, pp. 236-343, doi: 10.1145/3275116.3275150. [Online]. Available: http://dx.doi.org/10.1145/3275116.3275150

[165] P. Marti and I. Iacono, "Experience over time: evaluating the experience of use of a squeezable interface in the medium term," *Multimedia Tools and Applications,* vol. 76, no. 4, pp. 5095-5116, 2017, doi: 10.1007/s11042-016-3595-8.

[166] H. K. Kim, S. H. Han, J. Park, and W. Park, "How User Experience Changes over Time: A Case Study of Social Network Services," *Human Factors and Ergonomics in Manufacturing & Service Industries,* vol. 25, no. 6, pp. 659-673, Nov-Dec 2015, doi: 10.1002/hfm.20583.

[167] J. Hart and A. Sutcliffe, "Is it all about the Apps or the Device?: User experience and technology acceptance among iPad users," *International Journal of Human-Computer Studies,* vol. 130, pp. 93-112, Oct 2019, doi: 10.1016/j.ijhcs.2019.05.002.

[168] F. Sahar, J. Varsaluoma, S. Kujala, and K. Väänänen-Vainio-Mattila, "Identifying the user experience factors of a multi-component sports product," presented at the Proceedings of the 18th International Academic MindTrek Conference: Media Business, Management, Content & Services, Tampere, Finland, 2014. [Online]. Available: https://doi.org/10.1145/2676467.2676499.

[169] E. Karapanos, J. B. Martens, and M. Hassenzahl, "Reconstructing experiences with iScale," *International Journal of Human-Computer Studies,* vol. 70, no. 11, pp. 849-865, Nov 2012, doi: 10.1016/j.ijhcs.2012.06.004.

[170] J. Vissers, L. De Bot, and B. Zaman, "MemoLine: Evaluating long-term UX with children," in *12th International Conference on Interaction Design and Children, IDC 2013, June 24, 2013 - June 27, 2013*, New York, NY, United states, 2013: Association for Computing Machinery, in ACM International Conference Proceeding Series, pp. 285-288, doi: 10.1145/2485760.2485836. [Online]. Available: http://dx.doi.org/10.1145/2485760.2485836

[171] J. Varsaluoma and V. Kentta, "DrawUX: web-based research tool for long-term user experience evaluation," presented at the Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, Copenhagen, Denmark, 2012. [Online]. Available: https://doi.org/10.1145/2399016.2399138.

[172] M. Kurosu, A. Hashizume, Y. Ueno, T. Tomida, and H. Suzuki, "UX graph and ERM as tools for measuring Kansei experience," in *International Conference on Human-Computer Interaction*, 2016: Springer, pp. 331-339.

[173] M. Kurosu, A. Hashizume, and Y. Ueno, "User experience evaluation by ERM: Experience recollection method," in *20th International Conference on Human-Computer Interaction, HCI 2018, July 15, 2018 - July 20, 2018*, Las Vegas, NV, United states, 2018, vol. 10901 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 138-147, doi: 10.1007/978-3-319-91238-7_12. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-91238-7_12

[174] S. Kujala, T. Miron-Shatz, and J. J. Jokinen, "The Cross-Sequential Approach: A Short-Term Method for Studying Long-Term User Experience," *Journal of Usability Studies,* vol. 14, no. 2, pp. 105-116, Feb 2019. [Online]. Available: <Go to ISI>://WOS:000461836300005.

[175] C. V. L. T. Brennand, C. A. R. L. Brennand, V. R. M. L. Maike, J. V. da Silva, and M. C. C. Baranauskas, "Evaluating the User Experience: A Study on Childrens Interaction with Socio-enactive Artifacts in a Hospital Context," in *8th International Conference on Design, User Experience, and Usability, DUXU 2019, held as part of the 21st International Conference on Human-Computer Interaction, HCI International 2019, July 26, 2019 - July 31, 2019*, Orlando, FL, United states, 2019, vol. 11583 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 408-422, doi: 10.1007/978-3-030-23570-3_30. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-23570-3_30

[176] M. Burmester, M. Mast, K. Jäger, and H. Homans, "Valence method for formative evaluation of user experience," presented at the Proceedings of the 8th ACM Conference

on Designing Interactive Systems, Aarhus, Denmark, 2010. [Online]. Available: https://doi.org/10.1145/1858171.1858239.

[177] L. Colombo and M. Landoni, "A diary study of children's user experience with eBooks using flow theory as framework," in *13th International Conference on Interaction Design and Children, IDC 2014, June 17, 2014 - June 20, 2014*, Aarhus, Denmark, 2014: Association for Computing Machinery, in ACM International Conference Proceeding Series, pp. 135-144, doi: 10.1145/2593968.2593978. [Online]. Available: http://dx.doi.org/10.1145/2593968.2593978

[178] S. Kujala, T. Walsh, P. Nurkka, and M. Crisan, "Sentence completion for understanding users and evaluating user experience," *Interacting with Computers,* vol. 26, no. 3, pp. 238-255, 2014, doi: 10.1093/iwc/iwt036.

[179] C. Rusu, V. Z. Rusu, P. Munoz, V. Rusu, S. Roncagliolo, and D. Quinones, "On user experience in virtual museums," in *9th International Conference on Social Computing and Social Media, SCSM 2017 held as part of the 19th International Conference on Human-Computer Interaction, HCI International 2017, July 9, 2017 - July 14, 2017*, Vancouver, BC, Canada, 2017, vol. 10282 LNCS: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 127-136, doi: 10.1007/978-3-319-58559-8_12. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-58559-8_12

[180] B. Van De Laar, B. Reuderink, D. P.-O. Bos, and D. Heylen, "Evaluating user experience of actual and imagined movement in bci gaming," *International Journal of Gaming and Computer-Mediated Simulations,* vol. 2, no. 4, pp. 33-47, 2010, doi: 10.4018/jgcms.2010100103.

[181] I. Rodriguez, V. Herskovic, M. Karyda, and A. Lucero, "Exploring tangible ways to evaluate user experience for elders," in *2018 CHI Conference on Human Factors in Computing Systems, CHI EA 2018, April 21, 2018 - April 26, 2018*, Montreal, QC, Canada, 2018, vol. 2018-April: Association for Computing Machinery, in Conference on Human Factors in Computing Systems - Proceedings, p. ACM SIGCHI, doi: 10.1145/3170427.3188450. [Online]. Available: http://dx.doi.org/10.1145/3170427.3188450

[182] D. Raptis, A. Bruun, J. Kjeldskov, and M. B. Skov, "Converging coolness and investigating its relation to user experience," *Behaviour and Information Technology,* vol. 36, no. 4, pp. 333-350, 2017, doi: 10.1080/0144929X.2016.1232753.

[183] D. Raptis, E. Papachristos, J. Kjeldskov, M. B. Skov, and N. Avouris, "Studying the effect of perceived hedonic mobile device quality on user experience evaluations of mobile applications," *Behaviour and Information Technology,* vol. 33, no. 11, pp. 1168-1179, 2014, doi: 10.1080/0144929X.2013.848239.

[184] Z. Zhang, L. Zhang, and Y. Li, "The analysis of influence factors of user experience in bullet screen," in *AHFE Virtual Conference on Usability and User Experience, the Virtual Conference on Human Factors and Assistive Technology, the Virtual Conference on Human Factors and Wearable Technologies, and the Virtual Conference on Virtual Environments and Game Design, 2020, July 16, 2020 - July 20, 2020*, San Diego, CA, United states, 2020, vol. 1217 AISC: Springer, in Advances in Intelligent Systems and

Computing, pp. 272-278, doi: 10.1007/978-3-030-51828-8_36. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-51828-8_36

[185] S. Dittrich, F. Hof, and A. Wiethoff, "Interacdiff," in *2019 Conference on Mensch und Computer, MuC 2019, September 8, 2019 - September 11, 2019*, Hamburg, Germany, 2019: Association for Computing Machinery, in ACM International Conference Proceeding Series, pp. 583-587, doi: 10.1145/3340764.3344463. [Online]. Available: http://dx.doi.org/10.1145/3340764.3344463

[186] C. Krüger, T. Kojić, L. Meier, S. Möller, and J. Voigt-Antons, "Development and Validation of Pictographic Scales for Rapid Assessment of Affective States in Virtual Reality," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, 26-28 May 2020 2020, pp. 1-6, doi: 10.1109/QoMEX48832.2020.9123100.

[187] G. S. N. Meedin and I. Perera, "Crowdsourcing towards User Experience evaluation: An intelligent user experience questionnaire (IUEQ)," in *2014 14th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 10-13 Dec. 2014 2014, pp. 67-73, doi: 10.1109/ICTER.2014.7083881.

[188] A. Gross and S. Bongartz, "Why do i like it? Investigating the product-specificity of user experience," in *7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, NordiCHI 2012, October 14, 2012 - October 17, 2012*, Copenhagen, Denmark, 2012: Association for Computing Machinery, in NordiCHI 2012: Making Sense Through Design - Proceedings of the 7th Nordic Conference on Human-Computer Interaction, pp. 322-330, doi: 10.1145/2399016.2399067. [Online]. Available: http://dx.doi.org/10.1145/2399016.2399067

[189] C. Lallemand and V. Koenig, "Measuring the Contextual Dimension of User Experience: Development of the User Experience Context Scale (UXCS)," in *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 2020, pp. 1-13.

[190] F. Lachner, M.-A. Nguyen, and A. Butz, "Culturally sensitive user interface design: a case study with German and Vietnamese users," presented at the Proceedings of the Second African Conference for Human Computer Interaction: Thriving Communities, Windhoek, Namibia, 2018. [Online]. Available: https://doi.org/10.1145/3283458.3283459.

[191] S. Liu, T. Liang, S. Shao, and J. Kong, "Evaluating Localized MOOCs: The Role of Culture on Interface Design and User Experience," *IEEE Access,* vol. 8, pp. 107927-107940, 2020, doi: 10.1109/ACCESS.2020.2986036.

[192] T. Walsh, P. Nurkka, and R. Walsh, "Cultural differences in smartphone user experience evaluation," presented at the Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia, Limassol, Cyprus, 2010. [Online]. Available: https://doi.org/10.1145/1899475.1899499.

[193] T. R. C. B. de Souza and J. L. Bernardes, "The Influences of Culture on User Experience," Cham, 2016: Springer International Publishing, in Cross-Cultural Design, pp. 43-52.

[194] G. Hofstede, *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications, 2001.

[195] D. Winter, A. Hinderks, M. Schrepp, and J. Thomaschewski, "Welche UX Faktoren sind für mein Produkt wichtig?," *Mensch und Computer 2017-Usability Professionals,* 2017.

[196] H. B. Santoso and M. Schrepp, "Importance of user experience aspects for different software product categories," in *5th International Conference on User Science and Engineering, i-USEr 2018, August 28, 2018 - August 30, 2018*, Puchong, Malaysia, 2018, vol. 886: Springer Verlag, in Communications in Computer and Information Science, pp. 231-241, doi: 10.1007/978-981-13-1628-9_21. [Online]. Available: http://dx.doi.org/10.1007/978-981-13-1628-9_21

[197] E. Mortazavi, P. Doyon-Poulin, D. Imbeau, M. Taraghi, and J.-M. Robert, "Systematic Literature Review on UX dimensions and UX subjective evaluation tools," 2021.

[198] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate data analysis*, 7th ed. Pearson, 2014.

[199] A. Field, *Discovering statistics using IBM SPSS statistics*, 4th ed. Sage, 2013.

[200] L. G. Grimm and P. R. Yarnold, *Reading and understanding MORE multivariate statistics*. American psychological association, 2000.

[201] P. W. Jordan, *Designing pleasurable products: An introduction to the new human factors*. CRC press, 2000.

[202] G. Vitols and Y. Vitols-Hirata, "Impact of Culture Dimensions Model on Cross-Cultural Website Development," in *ICEIS (2)*, 2018, pp. 540-546.

[203] Hofstede-Insights. https://hi.hofstede-insights.com/national-culture (accessed.

[204] A. Hinderks, D. Winter, M. Schrepp, and J. Thomaschewski, "Applicability of User Experience and Usability Questionnaires," *Journal of Universal Computer Science,* vol. 25, no. 13, pp. 1717-1735, 2019. [Online]. Available: <Go to ISI>://WOS:000508640800004.

[205] H. Li, "Enhancing user experience of enterprise systems for improved employee productivity: A first stage of case study," in *2nd International Conference on HCI in Business, HCIB 2015 Held as Part of 17th International Conference on Human-Computer Interaction, HCI International 2015, August 2, 2015 - August 7, 2015*, Los Angeles, CA, United states, 2015, vol. 9191: Springer Verlag, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 493-500, doi: 10.1007/978-3-319-20895-4_45. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-20895-4_45

[206] J. Yoon, C. Kim, and R. Kang, "Positive User Experience over Product Usage Life Cycle and the Influence of Demographic Factors," *International Journal of Design,* vol. 14, no. 2, pp. 85-102, Aug 2020. [Online]. Available: <Go to ISI>://WOS:000576790900006.

[207] M. Hassenzahl, S. Diefenbach, and A. Göritz, "Needs, affect, and interactive products–Facets of user experience," *Interacting with computers,* vol. 22, no. 5, pp. 353-362, 2010.

[208] H. Ryu and J. Kim, "Evaluation of User Experience of New Defense Medical Information System," *Healthcare Informatics Research,* vol. 25, no. 2, pp. 73-81, Apr 2019, doi: 10.4258/hir.2019.25.2.73.

[209] M. Koelle, S. Ananthanarayan, and S. Boll, "Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies," in *Proceedings of the 2020 CHI Conference*

on Human Factors in Computing Systems: Association for Computing Machinery, 2020, pp. 1–19.

[210] X. Sun and A. May, "A comparison of field-based and lab-based experiments to evaluate user experience of personalised mobile devices," *Adv. in Hum.-Comp. Int.,* vol. 2013, p. Article 2, 2013, doi: 10.1155/2013/619767.

[211] A. Kantosalo and S. Riihiaho, "Quantifying co-creative writing experiences," *Digital Creativity,* vol. 30, no. 1, pp. 23-38, Jan 2019, doi: 10.1080/14626268.2019.1575243.

[212] M. H. Phan, J. R. Keebler, and B. S. Chaparro, "The development and validation of the game user experience satisfaction scale (GUESS)," *Human factors,* vol. 58, no. 8, pp. 1217-1247, 2016.

[213] F.-L. Fu, R.-C. Su, and S.-C. Yu, "EGameFlow: A scale to measure learners' enjoyment of e-learning games," *Computers & Education,* vol. 52, no. 1, pp. 101-112, 2009.

[214] J. Pallant, *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*, 7th ed. Routledge, 2020.

[215] T. R. Hinkin, "A brief tutorial on the development of measures for use in survey questionnaires," *Organizational research methods,* vol. 1, no. 1, pp. 104-121, 1998.

[216] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 7th ed. Pearson, 2019.

[217] J. Wang and X. Wang, *Structural equation modeling: Applications using Mplus*, 2nd ed. John Wiley & Sons, 2020.

[218] L. t. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Structural Equation Modeling: A Multidisciplinary Journal,* vol. 6, no. 1, pp. 1-55, 1999/01/01 1999, doi: 10.1080/10705519909540118.

[219] D. Gefen and D. Straub, "A Practical Guide to Factorial Validity Using PLS-Graph: Tutorial and Annotated Example, Communications of the AIS," ed, 2005, pp. 91-109.

[220] M. Hassenzahl, "The interplay of beauty, goodness, and usability in interactive products," *Human–Computer Interaction,* vol. 19, no. 4, pp. 319-349, 2004.

[221] Statista. "Distribution of worldwide social media users in 2022." https://www.statista.com/markets/424/topic/540/social-media-user-generated-content/#statistic1 (accessed 16 August 2022.

[222] "Facebook Comes Under Fire After Whistleblower and Leaked Documents Reveal Negative Impact on Girls." https://www.everydayhealth.com/public-health/facebook-comes-under-fire-after-whistleblower-and-leaked-documents-reveal-negative-impact-on-young-girls/ (accessed August 17, 2022).

APPENDIX A SEARCH STRINGS OF DIFFERENT DATABASES

Engineering village

Search fields in Engineering village

Time Frame	Language	Type of document	Search fields
2010-2021	English	Journal article, Conference paper	KY: subject, title, abstract

(((((((((evaluation NEAR/2 UX) OR (evaluating NEAR/2 UX) OR (evaluation NEAR/2 "User experience") OR (evaluating NEAR/2 "User experience") OR (measurement NEAR/2 UX) OR (measuring NEAR/2 UX) OR (measurement NEAR/2 "user experience") OR (measuring NEAR/2 "user experience") OR (tool NEAR/2 UX) OR (tool NEAR/2 "user experience") OR (dimension NEAR/2 UX) OR (component NEAR/2 UX) OR (attribute NEAR/2 UX) OR (aspect NEAR/2 UX) OR (indicator NEAR/2 UX) OR (factor NEAR/ UX) OR (element NEAR/2 UX) OR (dimension NEAR/2 "user experience") OR (component NEAR/2 "user experience") OR (attribute NEAR/2 "user experience") OR (aspect NEAR/2 "user experience") OR (indicator NEAR/2 "user experience") OR (factor NEAR/2 "user experience") OR (element NEAR/2 "user experience") OR (tools NEAR/2 UX) OR (tools NEAR/2 "user experience") OR (dimensions NEAR/2 UX) OR (components NEAR/2 UX) OR (attributes NEAR/2 UX) OR (aspects NEAR/2 UX) OR (indicators NEAR/2 UX) OR (factors NEAR/ UX) OR (elements NEAR/2 UX) OR (dimensions NEAR/2 "user experience") OR (components NEAR/2 "user experience") OR (attributes NEAR/2 "user experience") OR (aspects NEAR/2 "user experience") OR (indicators NEAR/2 "user experience") OR (factors NEAR/2 "user experience") OR (elements NEAR/2 "user experience"))wn KY))) AND (({ca} OR {ja}) WN DT)))) AND ({english} WN LA))

IEEE Xplore

Search fields in IEEE Xplore

Time Frame	Language	Type of document	Search fields
2010-2021	-	Journal article, Conference paper, Early access articles	All

((evaluat* OR measur* OR tool OR dimension OR component OR attribute OR aspect OR indicator OR factor OR element) NEAR/2 ("user experience" OR UX))

Web of Science

Search fields in Web of Science

Time Frame	Language	Type of document	Search fields
2010-2021	English	Proceeding paper, Journal, Early access	TS: Title, abstract and keywords

((TS=(("user experience" OR UX) NEAR/2 (evaluat* OR measur* OR tool* OR dimension* OR component* OR attribute* OR aspect* OR indicator* OR factor* OR element*))))

ACM Digital Library

Search fields in ACM Digital Library

Time Frame	Language	Type of document	Search fields
2010-2021	-	Journal article, Conference paper	Abstract

[Abstract: recordabstract:] OR [Abstract: "ux evaluation"] OR [Abstract: "evaluating ux"] OR [Abstract: "evaluation of ux"] OR [Abstract: "evaluate ux"] OR [Abstract: "measuring ux"] OR [Abstract: "measurement of ux"] OR [Abstract: "ux measurement"] OR [Abstract: "measure ux"] OR [Abstract: "ux tool"] OR [Abstract: "tool of ux"] OR [Abstract: "user experience evaluation"] OR [Abstract: "evaluating user experience"] OR [Abstract: "evaluation of user experience"] OR [Abstract: "evaluate user experience"] OR [Abstract: "measuring user experience"] OR [Abstract: "measurement of user experience"] OR [Abstract: "user experience measurement"] OR [Abstract: "measure user experience"] OR [Abstract: "user experience tool"] OR [Abstract: "tool of user experience"] OR [Abstract: "ux dimension"] OR [Abstract: "ux aspect"] OR [Abstract: "ux attribute"] OR [Abstract: "ux indicator"] OR [Abstract: "ux component"] OR [Abstract: "ux factor"] OR [Abstract: "ux element"] OR [Abstract: "user experience dimension"] OR [Abstract:

"user experience aspect"] OR [Abstract: "user experience attribute"] OR [Abstract: "user experience indicator"] OR [Abstract: "user experience component"] OR [Abstract: "user experience factor"] OR [Abstract: "user experience element"] OR [Abstract: "evaluating the ux"] OR [Abstract: "evaluation of the ux"] OR [Abstract: "evaluate the ux"] OR [Abstract: "measuring the ux"] OR [Abstract: "measurement of the ux"] OR [Abstract: "measure the ux"] OR [Abstract: "tool of the ux"] OR [Abstract: "evaluating the user experience"] OR [Abstract: "evaluation of the user experience"] OR [Abstract: "evaluate the user experience"] OR [Abstract: "measuring the user experience"] OR [Abstract: "measurement of the user experience"] OR [Abstract: "measure the user experience"] OR [Abstract: "tool of the user experience"] OR [Abstract: "dimension of the user experience"] OR [Abstract: "aspect of the user experience"] OR [Abstract: "attribute of the user experience"] OR [Abstract: "indicator of the user experience"] OR [Abstract: "component of the user experience"] OR [Abstract: "factor of the user experience"] OR [Abstract: "element of the user experience"] OR [Abstract: "dimensions of the user experience"] OR [Abstract: "aspects of the user experience"] OR [Abstract: "attributes of the user experience"] OR [Abstract: "indicators of the user experience"] OR [Abstract: "components of the user experience"] OR [Abstract: "factors of the user experience"] OR [Abstract: "elements of the user experience"] OR [Abstract: "dimension of user experience"] OR [Abstract: " aspect of user experience"] OR [Abstract: " attribute of user experience"] OR [Abstract: "indicator of user experience"] OR [Abstract: " component of user experience"] OR [Abstract: "factor of user experience"] OR [Abstract: " element of user experience"] OR [Abstract: "dimensions of user experience"] OR [Abstract: " aspects of user experience"] OR [Abstract: " attributes of user experience"] OR [Abstract: "indicators of user experience"] OR [Abstract: " components of user experience"] OR [Abstract: "factors of user experience"] OR [Abstract: " elements of user experience"]

APPENDIX B UX DIMENSIONS COVERED IN DIFFERENT UX EVALUATION TOOLS

No	UX tools	Usability	Utility	Affect	Hedonic	Engagement	Social	Informational	Sensory	Impact on body	Judgment	Physical	Cognitive	Overall UX	Others	Total
1	MEEGA+	✓	-	-	✓	✓	✓	✓	✓	-	✓	-	✓	✓	✓	10
2	Laine (2016)	✓	-	✓	✓	✓	-	✓	✓	-	✓	✓	-	✓	-	9
3	UX in IVE	✓	✓	✓	✓	✓	-	-	✓	✓	✓	-	-	-	✓	9
4	Jang (2019)	✓	✓	-	✓	-	✓	✓	✓	-	-	✓	-	-	✓	8
5	Moizer (2019)	✓	-	✓	✓	✓	-	-	✓	-	-	-	✓	✓	✓	8
6	Abro (2019)	✓	✓	✓	✓	✓	✓	-	-	✓	-	-	-	-	-	7
7	GEQ	-	-	✓	-	✓	✓	-	✓	✓	-	-	-	✓	✓	7
8	GUESS	✓	-	-	✓	✓	✓	-	✓	-	✓	-	-	-	✓	7
9	Krawczyk (2017)	✓	✓	-	✓	-	✓	-	✓	-	-	-	✓	✓	-	7
10	meCUE	✓	✓	✓	✓	-	-	-	✓	-	✓	-	-	✓	-	7
11	Nielsen (2018)	✓	✓	✓	✓	-	-	-	✓	-	-	✓	-	-	✓	7
12	Park (2018)	✓	✓	-	✓	-	✓	✓	✓	-	-	✓	-	-	-	7
13	Park (2013)	✓	✓	-	✓	-	✓	✓	✓	-	-	✓	-	-	-	7
14	Buchner (2012)	✓	✓	✓	-	-	-	-	-	✓	-	-	-	✓	✓	6
15	EDUGXQ	✓	✓	-	✓	✓	-	-	-	-	-	-	✓	-	✓	6
16	EGameFlow	✓	-	-	-	✓	✓	-	-	✓	-	-	✓	-	✓	6
17	Orehovacki (2018)	✓	✓	-	✓	-	✓	-	✓	✓	-	-	-	-	-	6
18	Pretto (2020)	✓	-	-	✓	-	-	✓	✓	-	-	-	✓	-	✓	6
19	Roth (2016)	✓	-	✓	✓	✓	-	-	✓	-	-	-	-	-	✓	6
20	SASSI	✓	✓	✓	✓	-	-	-	-	✓	✓	-	-	-	-	6
21	UEQ+	✓	✓	-	✓	-	-	✓	✓	-	-	-	-	✓	-	6
22	VRLEQ	✓	-	✓	-	✓	-	-	-	✓	✓	-	-	-	✓	6
23	Zhuo (2013)	✓	✓	✓	✓	-	✓	-	-	-	✓	-	-	-	-	6
24	Aestimo	✓	✓	-	✓	-	-	-	✓	-	-	-	-	✓	-	5
25	AttrakDiff	✓	✓	-	✓	-	-	-	✓	-	-	-	-	✓	-	5
26	CEGEQ	✓	-	-	✓	-	-	-	✓	-	✓	-	-	-	✓	5
27	CSUQ	✓	✓	-	✓	-	-	✓	-	-	-	-	-	✓	-	5
28	Cuadros (2021)	✓	✓	-	✓	✓	-	-	-	-	-	-	✓	-	-	5
29	FASER LX	✓	✓	-	✓	-	✓	✓	-	-	-	-	-	-	-	5
30	Lachner (2016)	✓	✓	-	✓	-	-	✓	✓	-	-	-	-	-	-	5
31	Macedo (14)	✓	✓	-	✓	-	-	-	✓	-	-	✓	-	-	-	5
32	Ormeno (2019)	✓	-	✓	✓	✓	-	-	-	-	✓	-	-	-	-	5
33	PSSUQ	✓	✓	-	✓	-	-	✓	-	-	-	-	-	✓	-	5
34	QUIS	✓	-	-	✓	-	-	✓	✓	-	-	-	-	✓	-	5
35	Saun (2021)	-	✓	-	-	-	-	-	-	✓	✓	✓	-	✓	-	5
36	UEQ	✓	✓	-	✓	-	-	-	✓	-	-	-	-	✓	-	5
37	UES	✓	-	-	✓	✓	-	-	✓	-	-	✓	-	-	-	5
38	UXQ4RL	✓	✓	-	✓	✓	-	-	-	-	-	-	✓	-	-	5
39	Van de laar (2010)	✓	-	-	-	✓	-	-	-	✓	-	-	-	✓	✓	5
40	VRNQ	-	-	-	✓	✓	-	-	✓	✓	-	-	-	-	✓	5

No	UX tools	Usability	Utility	Affect	Hedonic	Engagement	Social	Informational	Sensory	Impact on body	Judgment	Physical	Cognitive	Overall UX	Others	Total
41	Chinese UX quest.	✓	✓	-	✓	-	✓	-	-	-	-	-	-	-	-	4
42	EVEQ-GP	✓	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	4
43	Flow4D16	✓	-	-	✓	✓	-	-	-	✓	-	-	-	-	-	4
44	FSS	✓	-	-	✓	✓	-	-	-	-	-	-	-	-	✓	4
45	IMI	-	✓	-	✓	-	-	-	-	✓	-	-	-	-	✓	4
46	IVRUX	✓	-	-	-	✓	-	-	-	✓	✓	-	-	-	-	4
47	MAX	✓	✓	✓	-	-	-	-	-	-	✓	-	-	-	-	4
48	MUX	✓	-	-	-	-	✓	✓	-	-	-	✓	-	-	-	4
49	PENS	✓	-	-	✓	✓	-	-	-	-	-	-	-	-	✓	4
50	PQ	✓	-	-	-	✓	-	-	✓	-	-	-	-	-	✓	4
51	Reaction cards	✓	✓	-	✓	-	-	-	-	-	-	-	-	✓	-	4
52	SGUS	✓	✓	-	-	✓	-	✓	-	-	-	-	-	-	-	4
53	Short UEQ	✓	✓	-	✓	-	-	-	-	-	-	-	-	✓	-	4
54	SUXES	✓	✓	-	✓	-	-	-		-	✓	-	-	-	-	4
55	UES-SF	✓	-	-	✓	✓	-	-	✓	-	-	-	-	-	-	4
56	USE	✓	✓	-	✓	-	-	-	-	-	✓	-	-	-	-	4
57	UX Curve	✓	✓	-	✓	-	-	-	-	-	-	-	-	✓	-	4
58	UXS	✓	✓	-	✓	-	-	-	-	-	-	-	-	✓	-	4
59	VOF	✓	-	-	-	-	-	✓	-	✓	-	✓	-	-	-	4
60	CAS	✓	-	-	✓	✓	-	-	-	-	-	-	-	-	-	3
61	Fun toolkit	✓	-	-	✓	-	-	-	-	-	✓	-	-	-	-	3
62	GodSpeed questionnaire	-	-	✓	✓	-	-	-	-	-	-	-	-	-	✓	3
63	HED/UT	✓	✓	-	✓	-	-	-	-	-	-	-	-	-	-	3
64	Honeycomb	✓	✓	-	✓	-	-	-	-	-	-	-	-	-	-	3
65	iScale	✓	✓	-	✓	-	-	-	-	-	-	-	-	-	-	3
66	ITC-SOPI	-	-	-	-	✓	-	-	-	✓	-	-	-	-	✓	3
67	ITVUX	-	-	✓	✓	-	-	-	✓	-	-	-	-	-	-	3
68	SUISQ-R	✓	-	-	-	-	-	-	-	-	✓	-	-	-	✓	1
69	UEQ-S	✓	✓	-	✓	-	-	-	-	-	-	-	-	-	-	3
70	UXSC	✓	✓	-	✓	-	-	-	-	-	-	-	-	-	-	3
71	Yang (2018)	✓	-	-	-	-	-	✓	✓	-	-	-	-	-	-	3
72	Yu (2019)	✓	-	-	-	-	-	-	-	✓	-	✓	-	-	-	3
73	3E, 3E*	-	-	✓	-	-	-	-	-	-	-	-	-	✓	-	2
74	INTUI	✓	-	-	✓	-	-	-	-	-	-	-	-	-	-	2
75	PGQ	-	-	-	-	-	-	-	-	✓	-	-	-	✓	-	2
76	SUS	✓	-	-	-	-	-	-	-	-	✓	-	-	-	-	2
77	DALI	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
78	Comfort/Discomfort	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
79	Emoti-SAM	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	1
80	FKS	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	1
81	FMS	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
82	Geneva emo. wheel	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	1
83	GEQ (Engagement)	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	1
84	HabuT instrument	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	1
85	I-PANAS-SF	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	1
86	Klein (2020)	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	1

No	UX tools	Usability	Utility	Affect	Hedonic	Engagement	Social	Informational	Sensory	Impact on body	Judgment	Physical	Cognitive	Overall UX	Others	Total
87	Korber (2013)	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	1
88	MSAQ	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
89	NASA-TLX	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
90	NPS	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	1
91	PANAS	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	1
92	PPA	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	1
93	PrEmo	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	1
94	SAM	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	1
95	SEA	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
96	SEI	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	1
97	SEQ	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	1
98	SMEQ	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
99	SSQ	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	1
100	SUS	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	1
101	TACTUX	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	1
102	UNeeQ	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	1
103	UX needs scale	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	1
104	VisAWI	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	1

APPENDIX C ETHICS CERTIFICATE (CER-2021-12-D)

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

CER-2021-12-D

CERTIFICAT D'APPROBATION ÉTHIQUE

Le Comité d'éthique de la recherche de Polytechnique Montréal, selon les procédures en vigueur, en vertu des documents qui lui ont été fournis, a examiné le projet de recherche suivant et conclu qu'il respecte les règles d'éthique énoncées dans sa Politique en matière d'éthique de la recherche avec des êtres humains.

Projet	
Titre du projet	**Comparison of the selection of UX dimensions between experts and end users** **CER-2021-12-D**
Étudiant requérant	**Ehsan Mortazavi**, Candidat au PhD, Département de mathématiques et de génie industriel
Sous la direction de:	**Jean-Marc Robert, professeur titulaire**, Département de mathématiques et de génie industriel, Polytechnique Montréal & **Daniel Imbeau, professeur titulaire**, Département de mathématiques et génie industriel, Polytechnique Montréal.

Financement	
Organisme	CRSNG
No de UBR	3100408
Programme	Programme de subventions à la découverte
No d'octroi:	RGPIN-2018-06733
Titre original de l'octroi:	Ergonomie prospective : méthodes et données empiriques
Chercheur principal:	Jean-Marc Robert (PolyMTL)

MODALITÉS D'APPLICATION

Toute modification importante qui pourrait être apportée au protocole expérimental doit être transmise au Comité avant sa mise en œuvre.

L'équipe de recherche doit informer le Comité de tout élément ou évènement imprévu pouvant avoir une incidence sur le bien-être ou l'intégrité des participant(e)s impliqué(e)s dans le projet de recherche ainsi que tout problème susceptible d'avoir une incidence sur les membres de l'équipe de recherche.

Selon les règles universitaires en vigueur, un suivi annuel est minimalement exigé pour maintenir la validité de la présente approbation éthique, et ce, jusqu'à la fin du projet. Le questionnaire de suivi est disponible sur la page web du Comité.

Date de délivrance : **31 août 2020**

Date de fin de validité : **1er octobre 2021**

Farida Cheriet, présidente
Comité d'éthique de la recherche
Polytechnique Montréal

Date du prochain suivi : **1er septembre 2021**

Comité d'éthique de la recherche avec des êtres humains
Tél.: 514 340-4711 poste : 3755
Fax : 514 340-4992
Courriel : ethique@polymtl.ca

Adresse postale
C.P. 6079, succ. Centre-Ville
Montréal (Québec) Canada H3C 3A7

Campus de l'Université de Montréal
2900, boul. Édouard-Montpetit
2500, chemin de Polytechnique
Montréal (Québec) Canada H3T1J4

APPENDIX D RANKING OF UX DIMENSIONS PER PRODUCT CATEGORY

UX dimension	Word processing	Spreadsheet	Messenger	Social Network	Video Conferencing	Web shop	News Portal	Booking System	Info Web page	Learning Platform	Programming tool	Image Processing	Online Banking	Video Portal	Game
Content quality	11	10	16	4	15	2	1	4	1	1	11	10	9	1	9
Customization	12	12	12	12	12	15	18	13	18	14	5	9	12	7	14
Perspicuity	4	5	3	7	6	4	7	7	8	5	7	5	10	6	8
Efficiency	3	2	6	13	3	5	5	2	5	3	2	4	3	8	18
Immersion	20	19	20	16	18	19	12	20	17	13	15	16	20	9	1
Intuitive usage	8	4	4	5	4	6	9	8	7	10	10	8	7	3	6
Usefulness	2	1	9	17	2	7	4	5	2	2	1	1	2	14	21
Novelty	18	17	18	21	21	18	20	19	20	18	18	17	16	16	10
Beauty	16	15	14	15	19	13	15	14	14	17	16	14	15	15	7
Identity	19	20	13	8	16	20	19	18	19	20	19	19	17	21	20
Controllability	7	6	10	9	7	10	11	9	10	12	4	3	8	10	4
Stimulation	15	16	17	14	20	16	13	17	16	11	17	15	18	4	2
Clarity	5	7	8	10	9	9	2	6	4	8	8	6	5	5	11
Loyalty	14	14	19	19	17	14	10	15	13	19	14	18	13	18	17
Trust	6	9	2	3	8	3	3	1	6	9	13	13	1	11	19
Value	9	11	21	18	11	8	8	10	9	7	9	11	11	13	13
Ease of use	1	3	1	2	1	1	6	3	3	4	6	2	4	2	5
Error tolerance	10	8	15	20	13	11	17	11	12	15	3	7	6	17	16
Sociability	21	21	5	1	5	21	21	21	21	21	21	21	21	20	12
Social acceptance	17	18	7	6	10	17	16	16	15	16	20	20	19	19	15
Self-satisfaction	13	13	11	11	14	12	14	12	11	6	12	12	14	12	3

(1: the most important, 21: the least important)

APPENDIX E MEAN IMPORTANCE RATING (BOTTOM) AND RANKING (TOP) OF UX DIMENSIONS FOR DIFFERENT PRODUCT CATEGORIES

UX dimension	Word processing	Spreadsheet	Messenger	Social Network	Video Conferencing	Web shop	News Portal	Booking System	Info Web page	Learning Platform	Programming tool	Image Processing	Online Banking	Video Portal	Game
Content quality	9	8	11	2	10	1	1	3	1	1	9	8	7	1	7
	5.26	5.5	4.85	6.05	4.66	6.24	6.5	6.41	6.69	6.69	5.26	5.24	6.1	6.48	5.31
Customization	10	10	8	8	9	12	14	10	14	12	4	7	10	6	11
	4.92	5.33	5.13	5.45	5.24	4.68	4.07	4.81	3.19	5.08	6.18	5.59	5.21	5.52	4.81
Perspicuity	3	4	2	4	4	3	6	6	7	4	5	4	8	5	6
	5.87	6.13	6.1	5.82	5.98	6.07	5.45	6.06	5.59	6.05	6.0	5.95	5.82	5.56	5.4
Efficiency	2	2	4	9	2	4	5	2	4	3	2	3	3	7	13
	6.04	6.4	5.9	5.34	6.12	5.98	5.55	6.53	5.94	6.28	6.62	6.13	6.46	5.5	4.52
Immersion	16	15	15	12	13	15	11	16	13	11	12	13	16	8	1
	3.04	3.15	4.38	5.18	3.90	3.73	4.59	2.94	3.19	5.1	3.88	4.1	1.82	5.45	6.29
Intuitive usage	7	3	3	3	3	5	8	7	6	8	8	6	5	2	4
	5.58	6.2	6.05	5.98	6.02	5.98	5.25	6.0	5.84	5.67	5.71	5.67	6.23	6.02	5.71
Usefulness	1	1	6	14	1	6	4	4	2	2	1	1	2	12	16
	6.11	6.53	5.43	5.09	6.29	5.98	5.68	6.31	6.5	6.49	6.65	6.41	6.49	5.08	3.12
Novelty	14	14	13	16	16	14	16	15	16	14	15	14	13	14	8
	3.21	3.45	4.5	4.7	3.1	3.88	3.2	3.31	2.78	4.15	3.21	4.03	2.82	4.47	5.29
Beauty	13	12	10	11	14	10	13	11	11	13	13	11	12	13	5
	3.87	3.75	4.9	5.23	3.83	4.95	4.18	4.38	3.91	4.28	3.62	4.59	3.49	4.92	5.6
Identity	15	16	9	5	11	16	15	14	15	16	16	16	14	16	15
	3.09	2.88	4.95	5.73	4.29	3.17	3.68	3.47	3.0	3.69	3.03	3.46	2.79	3.18	3.79
Controllability	6	5	7	6	5	9	10	8	9	10	3	2	6	9	3
	5.58	6.00	5.33	5.70	5.83	5.34	4.70	5.69	4.53	5.18	6.21	6.15	6.15	5.34	5.81
Stimulation	12	13	12	10	15	13	12	13	12	9	14	12	15	3	2
	3.89	3.50	4.78	5.32	3.63	4.44	4.55	3.50	3.59	5.28	3.56	4.49	2.69	5.69	6.26
Clarity	4	6	5	7	7	8	2	5	3	6	6	5	4	4	9
	5.85	5.95	5.45	5.55	5.73	5.56	5.95	6.16	6.22	5.95	5..79	5.95	6.44	5.58	5.14
Loyalty	11	11	14	15	12	11	9	12	10	15	11	15	11	15	12
	3.98	4.03	4.43	5.00	3.95	4.76	4.98	4.31	3.94	4.00	4.12	3.82	4.97	4.11	4.60
Trust	5	7	1	1	6	2	3	1	5	7	10	10	1	10	14
	5.72	5.90	6.13	6.20	5.78	6.15	5.95	6.53	5.88	5.92	4.59	4.85	6.59	5.16	4.40
Value	8	9	16	13	8	7	7	9	8	5	7	9	9	11	10
	5.45	5.43	4.10	5.09	5.27	5.63	5.34	5.34	4.88	5.97	5.79	5.41	5.23	5.10	4.86

APPENDIX F ETHICS CERTIFICATE (CER-211-47-D)

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

CER-2122-47-D

CERTIFICAT D'APPROBATION ÉTHIQUE

Le Comité d'éthique de la recherche de Polytechnique Montréal, selon les procédures en vigueur, en vertu des documents qui lui ont été fournis, a examiné le projet de recherche suivant et conclu qu'il respecte les règles d'éthique énoncées dans sa Politique en matière d'éthique de la recherche avec des êtres humains.

Projet	
Titre du projet	**Design and development of the social scale of an evaluation tool for the user experience with interactive products** **CER-2122-47-D**
Étudiant requérant	**Ehsan Mortazavi**, Candidat au PhD, Département de mathématiques et de génie industriel
Sous la direction de:	**Daniel Imbeau, professeur titulaire**, Département de mathématiques et de génie industriel, Polytechnique Montréal & **Jean-Marc Robert, professeur associé**, Département de mathématiques et de génie industriel, Polytechnique Montréal.
Avec la collaboration de:	Philippe Doyon-Poulin (Co-directeur, PolyMTL)

Financement	
Organisme	CRSNG
No de UBR	3100408
Programme	Subvention à la découverte / Bourse pour l'étudiant
No d'octroi:	RGPIN-2018-06733
Titre original de l'octroi:	Ergonomie prospective : méthodes et données empiriques
Chercheur principal:	Jean-Marc Robert (PolyMTL)

MODALITÉS D'APPLICATION

Toute modification importante qui pourrait être apportée au protocole expérimental doit être transmise au Comité avant sa mise en œuvre.

L'équipe de recherche doit informer le Comité de tout élément ou évènement imprévu pouvant avoir une incidence sur le bien-être ou l'intégrité des participant(e)s impliqué(e)s dans le projet de recherche ainsi que tout problème susceptible d'avoir une incidence sur les membres de l'équipe de recherche.

Selon les règles universitaires en vigueur, un suivi annuel est minimalement exigé pour maintenir la validité de la présente approbation éthique, et ce, jusqu'à la fin du projet. Le questionnaire de suivi est disponible sur la page web du Comité.

Farida Cheriet, présidente
Comité d'éthique de la recherche
Polytechnique Montréal

Date de délivrance : **10 février 2022**

Date de fin de validité : **1er mars 2023**

Date du prochain suivi : **11 février 2023**

Comité d'éthique de la recherche
avec des êtres humains
Tél.: 514 340-4711 poste : 3755
Fax : 514 340-4992
Courriel : ethique@polymtl.ca

Adresse postale
C.P. 6079, succ. Centre-Ville
Montréal (Québec) Canada H3C 3A7

Campus de l'Université de Montréal
2900, boul. Édouard-Montpetit
2500, chemin de Polytechnique
Montréal (Québec) Canada H3T1J4