Old wine in new bottles or novel challenges? A critical analysis of empirical studies of User Experience

Conference Paper · May 2011

DOI: 10.1145/1978942.1979336 · Source: DBLP

CITATIONS
324

READS
418

2 authors:

Javier Andres Bargas-Avila
Google Inc.
35 PUBLICATIONS **1,415** CITATIONS

SEE PROFILE

Kasper Hornbæk
University of Copenhagen
186 PUBLICATIONS **5,694** CITATIONS

SEE PROFILE

Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience

Javier A. Bargas-Avila
University of Basel, Dept. of Psychology,
Cognitive Psychology & Methodology
Missionsstr. 62a, 4055 Basel, Switzerland
javier.bargas@unibas.ch

Kasper Hornbæk
Department of Computer Science,
University of Copenhagen
Njalsgade 128, 2300 Copenhagen, Denmark
kash@diku.dk

ABSTRACT

This paper reviews how empirical research on User Experience (UX) is conducted. It integrates products, dimensions of experience, and methodologies across a systematically selected sample of 51 publications from 2005-2009, reporting a total of 66 empirical studies. Results show a shift in the products and use contexts that are studied, from work towards leisure, from controlled tasks towards open use situations, and from desktop computing towards consumer products and art. Context of use and anticipated use, often named key factors of UX, are rarely researched. Emotions, enjoyment and aesthetics are the most frequently assessed dimensions. The methodologies used are mostly qualitative, and known from traditional usability studies, though constructive methods with unclear validity are being developed and used. Many studies use self-developed questionnaires without providing items or statistical validations. We discuss underexplored research questions and potential improvements of UX research.

Author Keywords

User experience, UX, Usability, HCI, Review, Meta-Analysis

ACM Classification Keywords

H.5.2 User Interfaces: Evaluation/methodology

INTRODUCTION

User Experience (UX) has emerged as an umbrella phrase for new ways of understanding and studying the quality-in-use of interactive products. Research on UX began around the turn of the millennium, initially employing labels such as pleasurable products [40], hedonic quality [26], or engineering joy [23]. Early UX research argues that existing usability research is too focused on task efficiency and work, and that more encompassing notions of quality are needed. Later research aims at defining, conceptualizing, and designing for UX (see [29] for a review).

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

In contrast to an instrumental, task-oriented view of interactive products, UX research focuses also on hedonic qualities of use. Such qualities concern for instance aesthetics [77] or self-actualization [20]. Another frequently mentioned characteristic is a focus on the positive emotions and affect that people experience while interacting with products [66, 29]. UX research focuses on the dynamics of experience, and on modeling how interactive products, person characteristics, and context work together in shaping the experience of use [88]. Finally, UX research calls for new methods and approaches for designing and evaluating experience.

The motivation for the present paper is twofold. First, many authors in the field of UX claim that they methodologically or content-wise break new ground. We would like to understand if and how this happens in empirical studies. Second, taking stock of empirical studies may help clarify what UX is; survey approaches defining UX have advanced the field, but left many open questions [50].

This paper aims to provide an integrated review of UX. We look for similarities across products, dimensions of experience, and methodologies. For researchers, we aim to provide an overview, to point out under-explored research areas and potential improvements of methodology. For practitioners, our paper serves as a synopsis of current research on UX.

RELATED WORK

The phrase *user experience* is used in several ways in the HCI literature. One is to denote the design and use of user interfaces, in effect working as a synonym for interaction, usability, or even user-centered design. Such use occurs in early [73] and recent [80] HCI papers. Another use of the phrase is to denote an emerging movement in research – a new paradigm – that focuses on non-instrumental needs and experience in a more complex sense [22, 29, 51, 50]; this understanding of UX forms the focus of the present paper. Our goal is to take stock of the state of the art in the scientific UX literature. We therefore do not discuss the many companies that work with UX, practical resources for UX design, or the abundant discussions on UX on websites and blogs.

The key focus of the UX movement is on the experience of interaction with interactive products. In contrast to usability research, UX research represents "a turn to experience" [91], and its motivation may be summarized as a search for "new approaches to the design of interactive products, which accommodate experiential qualities of technology use rather

than product qualities" [24, p. 353]. A substantial body of research has analyzed the nature of this experience (e.g., [12, 13]). The tenets that this focus entails are as follows:

1. UX takes a holistic view of users' interaction with interactive products. Most definitions emphasize that all aspects of product use are in focus, and some include anticipated use of products (e.g., [59]) and experiences following the use situation (e.g., [50]). Some authors stress that the holistic view also implies that users' experience is indivisible and needs to be reasoned about and studied as such [59]. Others put a particular emphasis on emotions: "emotion is at the heart of any human experience" and "emotion affects how we plan to interact with products, how we actually interact with products, and the perceptions and outcomes that surround those interactions" [12, p. 264]. The holistic perspective for some writers implies an attention to user experience in social contexts of use, leading to a focus on co-experience; that is, the creation of meaning and emotion among users through product use [12].

2. UX focuses on positive aspects of users' interaction with interactive products. Initial UX research departed from a stated dissatisfaction with removing usability problems and improving task completion time [26, 40]. Instead, it was proposed to focus on positive aspects of interaction, in particular, on hedonic, non-instrumental aspects. The term non-instrumental usually refers to task-unrelated qualities that fulfill general human needs [23], contrasting instrumental aspects, where utilitarian task completion is the main focus [32]. Non-instrumental aspects include for instance visual aesthetics [19, 64, 78] and beauty [21], joy of use [32], stimulation, personal growth, or surprise [29]. For many researchers, the emphasis on positive aspects of UX leads to a focus on human values and needs [24], because they ultimately determine why something is positive to users.

3. UX emphasizes the situational and dynamic aspects of using interactive products [29, 50] and the importance of context [50, 69]. Karapanos and colleagues [42] studied how experiences with a smart phone developed over several weeks, and how anticipation shaped subsequent experiences. Other researchers discuss that retrospective summary assessments do not reflect a whole experiential episode, but rather its most recent incidents [27]. They propose the usage of additional measurement strategies, such as repeated measurements throughout the experiential episode.

4. UX views and models the quality of interactive products as multidimensional. Research does not focus exclusively on the value of a product to accomplish tasks; it focuses also on symbolic and aesthetic value [54, 86]. For instance, Diefenbach and Hassenzahl [11] showed that the beauty of products is valued but discounted when it comes to choosing between a beautiful or a usable mobile phone; or van Schaik and Ling [86] explored the relation between usability, hedonic value, beauty, and goodness for user experience with web sites.

5. UX entails a need for new methods and approaches for designing and evaluating experience: "In interactive systems the challenge is to understand the influence small experiences and emotional responses have on others, as well as the overall view ... emotional responses are hard to understand, let alone quantify. New research methods are needed to better articulate the relationship between what we feel and what we do" [12, p. 265]. Many researchers regard traditional methods as being inapt for UX research: "A common strategy ... is the reduction of experience into a number of factors or processes ... such approaches may be useful for experimental analysis but they can miss some of the insights available in accounts that resist such reduction ... qualitative data provides a richness and detail that may be absent from quantitative measures." [76, p. 91-92].

Despite the progress in understanding the use of interactive products, the UX movement has also raised several questions, in particular regarding the products used in UX research, the dimensions of experience that are relevant for UX, and the methodology to be used. Next, we describe these questions as they provide the motivation for this review.

The *products* studied in UX research are debated: On the one hand some researchers posit that the focus of UX research should be about "everyday life" [3], on the other hand many appear willing to apply UX methods to any domain.

The *dimensions* of UX research are unclear. Law and colleagues [50] surveyed the views of 275 researchers and practitioners working with user experience. They found varying opinions on the definition of UX and its key characteristics. The relation of UX research to usability research is also contested. Some researchers write of UX as an "alternative to traditional HCI" [29, p. 91] and that it is not just "old wine in new bottles" [22, p. 11]; whereas it is elsewhere maintained that usability criteria can be used to assess aspects of user experience [36].

The *methodologies* applied are also debated and reignite the discussion between qualitative and quantitative approaches. A key issue is whether UX may be measured and modeled, and if so how: "A current trend seems to ascribe higher signicance to qualitative data analysis methods. Conversely, a danger would be that lower importance, priority and attention may be given to statistical methods." [51, p. 320].

In sum, although many papers discuss UX, its definition and distinct characteristics as a research field are currently unclear. Moreover, empirical findings are rarely synthesized. These limitations may be addressed normatively, by some vision of what UX research should be (e.g., [12, 29, 51]). They may also be answered empirically, by taking stock of the products, dimensions and methodologies of existing UX research; the present paper purports to do just that.

METHOD

The main goal of this study is to explore the products, dimensions, and methodological approaches used for UX research. To perform this analysis we decided to follow an analytical approach, using a representative sample of published empirical works in this field, covering a broad range of disciplines and approaches. The selection of publications was conducted in four steps, adapted from the QUOROM statement [62] which defines a procedure for conducting meta-analyses (see Figure 1).

Phase 1: Potentially relevant publications identified

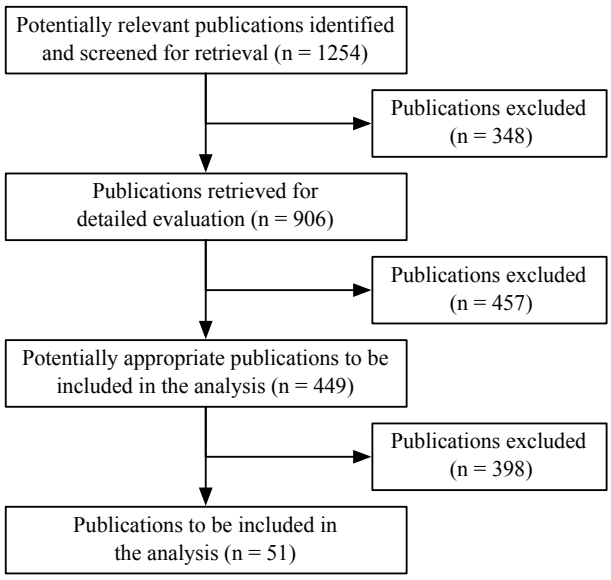
Source selection. A search in Google Scholar for "user experience" showed that relevant publications are spread across multiple scientific journals and conferences. Thus a procedure similar to [33], where a set of venues are defined and searched through, would not be appropriate, because many relevant sources may be missed. Instead we chose three scientific repositories spanning more than 13000 publication venues: ACM Digital Library (DL), ISI Web of Knowledge (WoK) and ScienceDirect (ScD). We restricted the search to a timeframe of five years (2005 to 2009).

Search term. We used the exact query "user experience". Although variations describe the same or very similar topics (such as customer experience, use experience) or subareas in this field (e.g., game experience, product experience), we found this term to best represent the new movement within HCI. Also, key publications in UX such as Forlizzi & Batterbee's framework [12] or Hassenzahl & Tractinsky's research agenda [29] use this expression. This procedure has several implications: (1) We miss publications that cover relevant topics for the field of UX, but refrain from using the term "user experience". This restriction seems acceptable, because our main goal is to review works of authors that intentionally aim to contribute to the UX movement. (2) Regarding false positives, we include publications that do not use the term "user experience" to denote a new paradigm (the oldest entry in DL using this term stems from 1970). False positives were excluded in later phases.

Search procedure. In DL and ScD "user experience" was used to search the publication's title, abstract or keywords. In the search interface of WoK such a restriction is not possible, so we used the phrase within "all fields".

Search results. The search returned in total 1254 entries (DL = 879, WoK = 270, ScD = 105) that were used for phase 2.

Figure 1. Adapted QUOROM procedure for this study.



Phase 2: Publications retrieved for detailed evaluation

First exclusion. All search results from phase 1 were imported into the software "Papers" and duplicates (items that describe identical publication) were removed (296). We also excluded entries with wrong years, that despite having narrowed the search to 2005-2009 were included (2). Next, all corresponding PDFs were downloaded. In this procedure four entries were excluded, because they did not refer to a publication in paper form. 46 entries were not accessible to our academic institutions. Thus, 906 publications remained.

Phase 3: Potentially appropriate publications

Screening criteria. In this phase the goal was to narrow the entries down to (1) original full papers that are (2) written in English and (3) speak in a broad sense about interactions between users and some kind of product(s). The last criterion was used because UX is usually seen as resulting from interactions with systems, products or services [50].

Second exclusion. According to the three criteria we excluded 299 entries because they contained no original full paper (e.g., workshops, posters, speeches, unrefereed magazine articles). Eight entries were not written in English and 150 were out of scope, treating topics like construction, software development, network protocols, or nursery. Note that these categories are not mutually exclusive – as exclusion criteria the first obvious category was chosen. After the second exclusion, 449 publications remained for phase 4.

Interrater reliability. The exclusion in phase 3 was made by the first author. To control for interrater effects, the second author performed the same categorization using 20% of the entries. The interrater reliability was found to be Kappa = 0.851 ($p < .001$), 95% *CI* (0.773, 0.929). According to [46], Kappa values of .8 and higher can be regarded as *outstanding*. The set was therefore used for phase 4.

Phase 4: Publications to be included in the analysis

Screening criteria. In the last phase we wanted to narrow the publication pool to studies that (1) contain empirical user data and (2) speak about "User Experience" as a new movement. The first criterion (empirical user data) required a publication to report qualitative or quantitative data collected from users. Case studies – even of a single user – were also included, as long as user data were reported in a clear way. Studies that were based on author's self-reflections or introspection, without involving users, were excluded. The second criterion (UX as a new movement) required that the authors mentioned UX as a new movement, an extension of traditional HCI, a new paradigm or more than just performance measures. Alternatively, the authors must state that they assess UX or explicitly mention that the constructs they assess are dimensions of UX and not equivalent to traditional usability measures. To make sure that the authors speak about UX as new movement, at least one of the following authors that we consider key to UX research had to be cited: Blythe, Forlizzi, Gaver, Hassenzahl, Lavie, Law, Mahlke, McCarthy, Monk, Tractinsky, Ward, Wright.

Third exclusion. According to the criteria above we excluded 136 entries that reported no empirical user data. An example

in this category is the UX research agenda [29]. We excluded 262 publications because they did not speak about UX as new movement (e.g., [79] where experience is researched as psychological perception), and/or did not mention one of the defined key authors (e.g., [16] where UX is used in the abstract but not defined as new movement nor are any key authors cited). Note that these categories are also not mutually exclusive. After the third exclusion a total of 51 publications remained for the final analysis.

Interrater reliability. The exclusion in phase 4 was also made by the first author. To control for interrater effects, the second author categorized 20% of the entries. The interrater reliability was again high, with Kappa = 0.821 ($p < .001$), 95% *CI* (0.652, 0.990).

A list with the 51 publications used for this review can be found on *http://bit.ly/gO2NQe*. For the analysis all descriptive data were entered into a database. Thus information such as results, measurements, measuring times, study types, study designs, participants, study setup, tasks, context, used products, was extracted for later anotation and analysis. Note that seven publications report two separate studies, and four report three. In the following sections, we often count and mark multiple studies as separate entities. In these cases, the number of studies sum up to N = 66.

RESULTS

In the following sections we report on products, dimensions and methodologies used in UX research.

Products, use situations and context in UX research

Table 1 shows an overview of the products in the sample. Among the most frequent products is art (21%). For instance, Blythe et al. [5] used an audio photography desk designed to challenge users' preconceptions about a city. Another example is van Boerdonk et al. [84], who developed an interactive canvas that aims to provide an alternative contact experience when meeting new people, while preventing visual prejudice.

The other products can be classified mostly as consumer or leisure products (e.g., mobile applications and phones, audio, video, TV applications). We categorized them into *leisure*, *work* or *mixed* (both leisure and work). In Table 1 it can be seen that leisure usage clearly dominates UX research (64%), whereas work or mixed are less frequent (each with 18%). It seems that the new movement of UX not only shifted the focus from usability to experience, but also changed the products researchers study.

Table 1. Products studied in UX research

Products used	e.g.	Total		Leisure		Work		Mixed	
		N	%	N	%	N	%	N	%
Art	[15]	14	21	14	100				
Mobile app/phone	[42]	14	21	5	36	4	29	5	36
Audio, video, TV	[55]	10	15	10	100				
Website	[86]	8	12			7	88	1	13
Imagined product	[31]	6	9	4	67			2	33
Interactive game	[56]	4	6	4	100				
No specific product	[67]	4	6	1	25			3	75
Other products		6	9	4	67	1	17	1	17

Notes. N = 66 studies; deviations from 100% due to rounding

Many researchers in the field of UX regard traditional measures such as task completion and performance as insufficient. We therefore looked at the use situations in which researchers study UX. We apply the term *use situation* to avoid possible confusion with tasks as used in usability research. We categorized them into three groups: (1) *Controlled tasks* stand for precisely defined assignments, which participants had to complete (e.g., find certain information on a website [86]); (2) *open use situations* stand for open instructions that explain the next steps, without providing detailed instructions or goals (e.g., instructions to play a game [56]); and (3) *user initiated use*, where users were free to choose if, when, and how they used the products (e.g., encourage users to use the product in their everyday life in tasks which they would be performing even when not evaluating a product [37]). The analysis shows that 61% of the studies use open use situations (n = 40), whereas 33% use controlled tasks (n = 22) and 20% leave the choice to participants (n = 13). Thus, UX focuses on more open use situations. Note that the total number of use situations surpass the number of studies, because some use multiple-use situation types.

Another frequently mentioned determinant of UX is *context* [29, 50]. We found that in almost half of the studies (45%, n = 30) researchers controlled the context by conducting their inquiry in a fixed setting. An example is [81], where UX of a multimodal media center is studied in a laboratory setting or [56], where UX of a game is measured in a psychophysiological lab. A third of the sample (33%, n = 22) used uncontrolled context and provided no descriptions of the physical or social setting. This is the case for most online studies conducted (e.g. [18, 11]). Only 21% (n = 14) described some parts of the context. This is often the case for studies using probes [39, 53] or methods similar to contextual inquiry [74, 70]. Considering that most researchers agree that UX is context dependent, it may be of interest to conduct more (field) studies exploring the role of context in UX.

Dimensions of Experience in UX research

A key question in UX is which experiential dimensions to assess. Table 2 shows all dimensions found in the sample.

The most frequent dimension is *Generic UX*. Here we summarize authors that do not further specify on which aspects of UX they focused. Common statements are "... it is possible to get a very rich view on the all-encompassing user experience" [44, p. 198] or "This paper reports on fieldwork, targeting user experiences of and with security technology" [57, p. 285]. *Generic UX* stems mostly from qualitative studies, where focus groups, interviews or probes are used (81% of *generic UX* dimensions originate from qualitative studies). In these cases, authors often do not report the dimensions that they focus on. Note that some of these studies do derive concrete conclusions about various product-specific dimensions of UX (e.g., [76, 53]).

The most frequent dimensions of UX that were assessed are *emotions and affect* (24%), followed by aspects of *enjoyment* (17%) and *aesthetics* (15%). This was unsurprising, as all three are often-mentioned core dimensions of UX. Although the dimension of aesthetics is often approached with experimental studies, the effect of emotion and enjoyment is

Table 2. Dimensions in UX research

UX dimension **	Examples and sources	N	%*
Generic UX	Interviews at the start and end of the trial [63]; Users made collages to express UX with mobile phones [61]	27	41
Affect, emotion ***	Affect measured with SAM [47] scale [28]; Emotional responses measured with psychophysiology [56]	16	24
Enjoyment, fun	Rate 13 play categories on a 5-point scale [8]; Pictures drawn after game experiences were coded e.g. for "fun" [92]	11	17
Aesthetics, appeal	Classical and expressive aesthetics [49] measured [87]; Website attractiveness [75] measured [19]	10	15
Hedonic quality	AttrakDiff [21] used to measure hedonic quality of websites [86]	9	14
Engagement, flow	FSS [38] used to measure flow [15]; Semi-structured interviews to understand engagement with technology [67]	8	12
Motivation	Probes used to understand motivation behind selecting a specific home [39]	5	8
Enchantment	Interview with a person to understand enchantment of technology [65]	4	6
Frustration	When people expressed aspects that they did not like (in semi-structured interviews) [4]	3	5
Other constructs	Values, spontaneity, ...	15	23

Notes. N = 66 studies * does not sum up to 100% because studies can measure several dimensions ** If authors mention exploring a certain dimension, it was counted as such without judging if the work really reports it *** We categorized *affect* and *emotion* in the same group, as they are treated mostly as synonyms in UX research. Many researchers in psychology, however, consider affect and emotions to be different (e.g., [72])

Table 3. UX data collection methods

Collection method	Examples and sources	N	%*
Questionnaires	Affect measured with SAM scale [30]; user feedback assessed with self-developed questionnaire [4]	35	53
Interviews (semi-structured)	Interview regarding interaction experience [84]; interview to elicit reactions regarding engagement [67]	13	20
User observation (live)	In-situ observation of PDA usage [70]; observation of people experiencing an artistic installation [8]	11	17
Videorecordings	Recordings of interactions with canvas [84]; videos to capture listening experiences on the move [52]	11	17
Focus groups	Focus group at end of trial to reflect PAD usage [70] ; group discussion to investigate preferences [19]	10	15
Interviews (open)	Interview to understand enchantment [65]; interviews regarding photo sharing experience [63]	8	12
Diaries	Emotions assessed with diaries [37]; diaries using day reconstruction & experience narration [42]	7	11
Probes	Participants were given a probe kit with a brief personal explanation and instruction [53]	6	9
Collage or drawings	From a list of topic areas participants chose an experience to articulate through collage [61]	5	8
Photographs	Encouraged users to take photographs related to attributes in their domestic environment [39]	5	8
Body movements	Choreography of interaction with canvas was evaluated by analysing the movements [84]	3	5
Psychophysiological measures	Psychopyhsiology (galvanic skin response, EMG, heart rate) of a game experience recorded [56]	3	5
Other methods	Think aloud, personal meaning maps, ...	18	27

Notes. N = 66 studies * data do not sum up to 100% because studies can use more than one method

never subject to experimental manipulation. An additional 14% of the studies focus on hedonic quality – all of them using AttrakDiff [21] or a modified version thereof (e.g., [83]).

We found many authors that propose new dimensions for understanding UX. These dimensions include enchantment [65], engagement [67], tangible magic [92], aesthetics of interaction [89], and relevance [76]. We will describe this phenomenon in the discussion section.

In addition, we looked at how many of these dimensions are usually assessed. Almost half of the studies (45%, n = 30) assessed only one, and 71% (n = 45) two or fewer dimensions. Half of the studies where only one dimension is assessed (n = 15) are categorized as "Generic UX"; in the other half, researchers focused on only one dimension (e.g., [65], where enchantment is the focus). Conversely, only 6% cover four or more dimensions (e.g., [56], where challenge, emotion, engagement, enjoyment and frustration are measured). It seems interesting that UX – despite being thought of as a multifaceted construct – is often approached in quite selective and restricted ways.

For Table 2, all assessments of usability were removed, even if the authors report them as being part of UX. A total of 45% (n = 30) of the studies also assessed usability metrics alongside UX (e.g., efficiency, effectiveness, satisfaction). Half of the studies that assessed both UX and usability (n = 15) looked into the relation between these two constructs (e.g., [55, 86]), whereas the other half refrained from doing so (e.g., [81, 90]). Studies with no usability metrics employed mainly qualitative or mixed study designs (90%, n = 27).

How are data on UX collected?

Half of the studies are qualitative (50%, n = 33), whereas 33% (n = 22) use quantitative methods, and 17% (n = 11) use both methodological approaches. These figures differ considerably from traditional HCI. Barkhuus and Rode analyzed 24 years of CHI publications [2]. For the year 2006 (the only reported year that falls within our sample), they found that only 14% of the studies were qualitative, whereas about 65% were quantitative (the rest used mixed designs or were not empirical). Thus the shift toward UX seems to be accompanied by a change in methodology.

Table 3 lists all UX data-collection methods found. Questionnaires are the dominant UX assessment method (53%). Others apply qualitative methods from HCI that involve talking directly with users, such as semi-structured interviews (20%), focus groups (15%) and open interviews (12%), as well as less intensive procedures such as user observation (17%), analysis of videorecordings (17%) and diaries (11%). We found an emerging group of constructive or projective methods such as probes, collage/drawings, or photographs. Objective measurement of UX via psychophysiology is used rarely. Thus data on UX are currently mostly collected with methods borrowed from traditional HCI.

More than half of the publications used questionnaires as a way of assessing UX. Table 4 shows an analysis of the surveys used for these 35 studies. Half of them (51%) use self-developed questionnaires but do not provide readers with the items used. Blythe et al., for example, conduct an online study with 563 users to analyze their experiences us-

ing a location-sensitive interactive play. The authors provide some selected conclusions, but refrain from reporting the content of the questionnaire [4]. A further example is Lankes et al., where a "standardized questionnaire, consisting of thirteen questions on the user experience factors fun, engagement and cognitive load" is used without providing any details about the items or the source [48, p. 255]. Another 29% also use selfmade questions, but provide them in the text (e.g., O'Brien et al. [67], where all items and their intended use are listed in the appendix), while 11% report them only partially (e.g., van Boerdonk et al. [84], where some selected items are reported as examples in the text). The finding that many researchers use selfmade items without providing them is not new and was also found for subjective satisfaction metrics in usability [33]. In terms of transparent study reporting, researchers as well as reviewers should aim at changing this in the future.

Table 4. Questionnaires in UX research

Used questionnaire	N	%*
Self-developed (items unknown)	18	51
Self-developed (items listed in the paper)	10	29
AttrakDiff [21]	7	20
Lavie & Traktinsky [49]	7	20
SAM [47]	7	20
Self-developed (items partly known)	4	11
Other surveys (e.g., FSS [38], IMI [58], Emocards [10])	11	35

Notes. n = 35 studies that used questionnaires * data do not sum up to 100% because some studies use more than one questionnaire

The most frequently used validated instruments cover Hassenzahl's hedonic quality [21] and Lavie & Tractinsky's aesthetics [49] with 20% each – a finding that is in line with frequently assessed UX dimensions (see Table 2). The most frequent tool for measuring emotions is Lang's SAM scale [47].

Note that 6 of the 11 studies that assessed "enjoyment" used self-developed items; not a single validated questionnaire was found in the sample. Remember that this dimension is one of the most frequently assessed in UX research (see Table 2). The reason why researchers who choose to use questionnaires stick to ad-hoc scales and refrain from using validated enjoyment scales remain unclear. There are validated scales with only 3 to 4 items (e.g., [85, 35]) and also instruments that were developed to measure more complex constructs like "Cognitive Absorption", where "heightened enjoyment" is assessed [1]. These are rich ressources mostly untapped by UX researchers.

When is UX assessed?

UX occurs before, during and after interaction [50]. Therefore we analyzed at which points in a study researchers chose to assess UX. We categorized the measuring points into *before*, *during*, and *after* interaction, as well as a fourth category labeled *imagined interaction*. In the latter studies, researchers did not look at real product interactions triggered by the study set-up, but used fictive products or retrospective techniques to analyze experiences from the past. Table 5 gives an overview of the temporal aspects of assessment.

Data show that frequencies increase over time, with *before* measurements being rare (20%) and *after* measurements being the most frequent (70%). Correspondingly, the most frequent pattern is the combination of *during* and *after* measurements – similar to traditional usability metrics, where users are observed when interacting and satisfaction is measured afterwards. Only 17% of the studies measure *before*, *during*, and *after* interaction.

Table 5. Temporal aspects of UX assessment

	before	during	after	imagined	N	%	Examples
		x	x		19	29	[39, 56]
			x		14	21	[92, 45]
	x	x	x		11	17	[76, 42]
		x			8	12	[70, 57]
	x		x		2	3	[81, 5]
				x	12	18	[67, 31]
N	13	38	46	12			
%*	20	58	70	18			

Notes. N = 66 studies * data do not sum up to 100% because studies can use multiple measurement times

Anticipated use assumes an important role in the field of UX, forms part of the ISO 9241-210 definition [36], and is a major difference to traditional HCI regarding temporal aspects. We therefore took a closer look at what researchers measure before interaction occurs. Thirteen studies used *before* assessments (see Table 5), of which only *expectations regarding the studied products* (36%, n = 5) was related to anticipated use. An example for this is the study of Turunen et al., where participants were asked about their expectations regarding multimodal media center applications [81]. The other assessed aspects were *prior experiences* with a similar product and *relevant habits* for the study (each with 36%, n = 5), as well as *open questions*, *attitudes* and *affective states* (each with 7%, n = 1). Thus, it seems that before measurements in UX are still widely disregarded.

The analysis of temporal aspects shows an additional phenomenon of current UX research: There are no truly longitudinal studies published. Some papers study experience over several weeks (e.g., [42, 52]), but projects that cover typical product life cycles of several months or even years are missing.

In addition to the reported analysis we grouped the papers by their main goal. This led to four groups: (1) *Basic research* (29%, n = 15) contains publications that investigate specific dimensions, the dynamics or the influencing factors to understand how UX works. (2) *Evaluation* (31%, n = 16) covers authors who developed a certain product or interface and conducted a study to evaluate its UX, such as an interactive canvas [84], a photo sharing service [63] or a tangible audio museum guide [89]. (3) *Methods* (29%, n = 15) consists of studies where a new method to design for, improve, measure, or explore UX is developed or applied. Finally a smaller group of publications covered (4) *requirements* (10%, n = 5), where authors aim at eliciting requirements for future products. In the following sections we will report on the groups *basic research*, *methods*, and *requirements*. Note that the papers in the group *evaluation* are too heterogenous to be discussed in detail.

Basic research

The group labeled as basic research contains publications that describe how UX is created, its core dimensions, and its development over time. There are two major subgroups consisting of four publications each: The first looks at the dimension of aesthetics, and its relation to usability [55, 86, 19, 87]. The other investigates how product judgements and choices are made and the factors influencing them [30, 18, 28, 11]. The remaining papers look at various dimensions of UX (e.g., engagement [67] or enchantment [65, 71]) or research how UX develops over time [42]. Surprisingly the authors within this group often do not relate to each other's work. Papers on basic research reference each other only six times, of which four are self-citations. It seems that for the time being, basic research in UX is quite heterogenous, going in various often unrelated directions.

Methods

Fifteen papers were mainly about the methodology of UX evaluation. The most prominent difference to mainstream HCI methods is that several papers present what may be called constructive or projective evaluation methods. Such methods ask participants to create one or more artifacts and use those to understand or assess UX. In four papers, the artifacts are visual representations, such as drawings [92] and collages [61]. Xu and colleagues [92] investigated how drawings may be used for interpreting children's experiences with technology. They developed a scheme for coding the drawings and showed how codes could differentiate three dimensions of UX: Fun, Goal Fit, and Tangible Magic. Others used video as the artifact produced to reflect experiences (e.g., [37]). Another constructive technique studied was Cultural Probes [14]. Lucero et al. [53] described how probes benefitted the design of new technologies for bathrooms, highlighting how probes help identify unexpected uses of the bathroom and look into participants' everyday lives.

Several papers presented evaluation techniques that did not rely on construction. These include papers using physiological measures of affect. For instance, Mandryk et al. [56] investigated heart rate, galvanic skin response and other measures, and showed how they could be used to distinguish gaming against a computer from against a friend. Others addressed UX methodology in ways resembling usability research: [82] described an expert evaluation method aimed at tapping UX and [83] adapted AttrakDiff to a work context.

The papers concerning methodology raise two open questions. First, papers in our sample rarely compare the new methods they propose to existing methods, or to alternative ways of gathering data on user experience (e.g., [7, 92, 53, 83], though see [37] for an exception). Thus, it is currently unknown whether the results obtained are valid or whether we could have, for instance, achieved the results obtained by probing in [53] by simply interviewing participants. Second, papers on constructive evaluation methods repeatedly discuss the difficulty of interpretation and data analysis. For instance, Xu et al. point out that the coding scheme is "open to evaluator's interpretation" [92, p. 224] and [37] discuss how interpretation of a diary method for expressing emotions and experiences is highly challenging.

Expectations and requirements

A considerable amount of studies (15%) report *expectations* or *requirements* for current or future products. Note that some researchers even report requirements and expectations within the same study and regard them as describing different things (e.g., [57]). Although it is important to look at what users expect from products, we would not consider these topics to be a dimension of UX (hence we did not code them in Table 2).

In most studies of expectations, users could not interact with a product – the research targets users' needs and imagined applications of these products (e.g., [74, 31]). Some publications treat requirements and expectations as one of many topics (e.g., Karapanos et al. [42] where expecations regarding the iPhone are treated), while others use it as main focus (e.g., Pyykkö et al. [41], where requirements regarding mobile 3D TV are elicited).

DISCUSSION

Next, we discuss what our findings mean to UX research, what to do (and perhaps do differently), and address the questions posed in the Related Work section.

Products, use situations, and goals

Our review suggests a heavy emphasis on art and consumer products. The phenomenon that art is often used in UX research is interesting, because the UX definition by Law et al. positions *Art* as being out of scope for UX [50, Figure 2]. Apparently the UX community disagrees on this point.

Further the turn to art and consumer products may preclude conclusions about experience with other types of products. The broadening of products studied is one of the purposes of UX research; it is fruitful, enriching, and an interesting expansion in scope. At the same time we find many studies in our sample that speak generally about UX, yet fail to reflect on the products used to study UX. The UX movement criticized traditional HCI for focussing only on work related products. Correspondingly we think that a narrow focus on consumer products and art only is comparably harmful to UX research, because the contribution to the understanding of a broad amplitude of products is largely ignored.

One common understanding is that non-instrumental, hedonic or non-task-oriented goals are associated with UX, whereas instrumental, pragmatic or task-oriented goals are associated with usability [29]. Initially, we aimed at categorizing the use situations into these two groups, but this proved difficult: Instrumental and non-instrumental goals were often interwoven and inseparable. For instance, one study even reported that users – not having received instrumental goals – started to invent such goals: "Although the designers did not deliberately structure goals or objectives for users, some imposed their own goals." [4, p. 132]. Recently, Law and van Schaik argued that the idea of applying type of goal as a differentiator between usability and UX is questionable [51]. Our data confirm this argument: While these concepts are certainly useful to reason about the influence of goals, they are not suited to serve as differentiators between traditional HCI and UX.

Dimensions of experience

As expected, affect and aesthetics are often used as indicators for quality of UX. With respect to affect, the most frequently used instrument is Lang's SAM scale [47]. With this scale, users provide ratings of emotion by selecting versions of a manikin that symbolizes different emotional states. This scale is well validated and widely used, but it is a relatively simplistic form of measuring emotions. We find it interesting that not a single UX study seeks to obtain more detailed measures of emotions, while at the same time many researchers agree that emotion is one of the key factors of UX [50]. In emotional psychology there are many established and validated ways of measuring emotions that provide more detailed and richer data – for an overview see [6]. Future research might benefit from these to do in depth studies of affective states in UX.

With respect to aesthetics, validated questionnaires are used [49, 21]. Recently, however, it was suggested that these questionnaires may be similar on subscales: "We argue that expressive aesthetics and hedonic quality are strongly overlapping constructs" [25, p. 254]. This phenomenon probably originates in the parallel development and publishing of these constructs and shows that a future consolidation may be beneficial to the field.

The studies reviewed also proposed many new dimensions of quality or variations of already known dimensions. We already mentioned enchantment [65], engagement [67], tangible magic [92], aesthetics of interaction [89], and relevance [76]. This variety in dimensions is inspiring to the field and supports new ways of thinking about experience. The main problem of this dimensionality explosion is that the relation to established constructs is rarely made clear. Let us briefly exemplify the problem using the notion of "enchantment". According to McCarthy and colleagues, enchantment is a "useful concept to facilitate closer relationships between people and technology" [60, p. 369], derived from a discussion of experiences of film and cell phones and used in many subsequent studies (e.g., [71, 65]). Although the notion of enchantment may be useful in many ways, one cannot help ponder how it relates to other constructs: Is enchantment a crucial and distinct component of UX? What is the difference between enchantment and established concepts such as for instance the experience of flow [9]? Is flow a condition for enchantment? Or a consequence? No studies in our sample report how enchantment can be measured and consequently help clarify its role within UX and the difference to other concepts. We argue that if new dimensions are needed and proposed, they should at some stage be accompanied by studies that clarify their difference to established constructs (see e.g. [68], where this step is partly taken for the dimension "engagement"). Else we risk ending up with an endless number of words that describe similar phenomena within UX, something from which the characterization of satisfaction in usability research has suffered [33].

Methodology

We have shown how research in UX draws on multiple methods. Although this in theory is attractive, in practice it leads to dichotomic research. On the one hand, some papers study very particular use situations, emphasizing richness of description and using mainly qualitative research methods (we call those *uniqueness studies*, e.g., [57, 65]). On the other hand, some studies model the dimensions of UX, emphasizing findings that generalize and using mostly quantitative research methods (we call those *dimension studies*, e.g., [11, 30, 87]). Settling the debate about the relative merits of these types of study is outside the scope of this review. Nevertheless, we argue that some studies overemphasize their methodological stance to the extent of damaging research quality. Many uniqueness papers do not report interview questions or protocols, rarely describe data analysis methods, focus mostly on generic UX, and contribute to the dimensionality explosion mentioned above. Many dimension papers do not attempt to study complex, ongoing interaction (often using screenshots or studying interaction lasting less than 30 minutes), and some say surprisingly little about experience and self-reports on UX. Unfortunately, few studies combine the methods of uniqueness and dimension studies (though see [42]) and our review shows a sad lack of reference between the groups.

The papers that contribute to method raise three points. First, the proposals for new methods are rarely validated. In work on usability evaluation, many papers have highlighted the difficulty of comparing methods [17, 34]. In work on UX methods, these difficulties receive scant attention: New methods are merely used without comparison to other methods, or the comparisons are weak. We see much opportunity here to improve our understanding of the relative merits of methods aimed at assessing or evaluating UX. Second, what we call constructive methods (e.g., sketches, probes) are not only rarely validated, they raise many issues about the process and validity of interpreting their results. Many authors use these methods as inspiration for design (as it may be argued is the purpose of cultural probes), but not all treat the methods this way. Third, some authors argue the need to rely on first-person methods for understanding experience. In contrast, much work on user centered design and usability research emphasize that we need to look at behavior, what people do, rather than merely listen to what they say and what they say they do. The tension between these approaches needs to be understood much better.

Limitations and future work

The aim of the present paper has been to critically discuss empirical studies of UX. Our approach has at least three shortcomings. First, we have focused on research studies and their treatment of UX, ignoring how practitioners work with UX. Although many of the publications we reviewed are from commercial research labs, additional work should review practice. Second, we have favored representativeness of studies through careful sampling. Although this has allowed us to broadly characterize the empirical studies that the UX movement has inspired, it means that the discussion of individual studies has been brief. Third, we have disregarded studies that explore UX through design, without empirical data. A review of design research could give an interesting and complementary view on UX research.

Our work has a number of implications for research in UX that should be clear from the above. Let us mention four additional areas in which more work is needed. First, rich descriptions of context were surprisingly absent in the studies reviewed. Even papers supposedly strong on capturing context reported not being able to reason about context sufficiently well [57]. Second, the basic research characterizing experience over time is meagre. Karapanos et al. [42] provided important insights, but their study nevertheless did not look at individual differences in use and the phases of use proposed were not clearly related to progression in time, neither for the individual participant nor across participants. Third, although multidimensionality is a key tenet of UX, surprisingly few studies describe several aspects of UX and their interrelation. Fourth, UX on desktop computers with productivity applications are rarely researched.

In sum, we have shown that UX research in several ways builds on existing research on HCI and usability. At the same time, the papers reviewed have helped identify novel challenges. We hope that our review helps take up and further develop existing solutions, while also addressing the distinct and pressing challenges in UX research.

REFERENCES

1. R. Agarwal and E. Karahanna. Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly*, 24(4):665–694, 2000.
2. L. Barkhuus and J. Rode. From Mice to Men–24 years of Evaluation in CHI. *Proc. CHI '07*, 2007.
3. K. Battarbee and I. Koskinen. Co-experience: user experience as interaction. *CoDesign*, 1(1):5–18, 2005.
4. M. Blythe, J. Reid, P. Wright, and E. Geelhoed. Interdisciplinary criticism: analysing the experience of riot! *Behaviour & Information Technology*, 25(2):127–139, 2006 *.
5. M. Blythe, J. Robinson, and D. Frohlich. Interaction design and the critics: what to make of the weegie. In *Proc. NordiCHI '08*, 53–62. ACM, 2008 *.
6. J. Coan and J. Allen. *Handbook of emotion elicitation and assessment*. Oxford University Press, 2007.
7. B. Costello and E. Edmonds. A study in play, pleasure and interaction design. In *Proc. DPPI '07*, 76–91. ACM, 2007 *.
8. B. Costello and E. Edmonds. Directed and emergent play. In *Proc. C&C 09*, 107–116. ACM, 2009 *.
9. M. Csikszentmihalyi. Flow and the Psychology of Discovery and Invention. *HarperPerennial, New York*, 1997.
10. P. Desmet. *Designing Emotions*. PhD thesis, Delft University of Technology, 2002.
11. S. Diefenbach and M. Hassenzahl. The Beauty Dilemma: beauty is valued but discounted in product choice. In *Proc. CHI 2009*, 1419–1426. ACM, 2009 *.
12. J. Forlizzi and K. Battarbee. Understanding experience in interactive systems. In *Proc. DIS '04*, 261–268. ACM, 2004.
13. J. Forlizzi and S. Ford. The building blocks of experience: an early framework for interaction designers. In *Proc. DIS '00*, 419–423. ACM, 2000.
14. W. Gaver, T. Dunne, and E. Pacenti. Cultural probes. *Interactions*, 6(1):21–29, 1999.
15. S. Gilroy, M. Cavazza, and M. Benayoun. Using affective trajectories to describe states of flow in interactive art. In *Proc. ACE '09*, 165–172. ACM, 2009 *.
16. L. Gong. Is happy better than sad even if they are both non-adaptive? Effects of emotional expressions of talking-head interface agents. *Int. Journal of Human-Computer Studies*, 65(3):183–191, 2007.
17. W. Gray and M. Salzman. Damaged merchandise? A review of experiments that compare usability evaluation methods. *Human-Computer Interaction*, 13(3):203–261, 1998.
18. J. Hartmann, A. De Angeli, and A. Sutcliffe. Framing the user experience: information biases on website quality judgement. In *Proc. CHI '08*, 855–864. ACM, 2008 *.
19. J. Hartmann, A. Sutcliffe, and A. Angeli. Towards a theory of user judgment of aesthetics and user interface quality. *ACM TOCHI*, 15(4):1–30, 2008 *.
20. M. Hassenzahl. The thing and I: understanding the relationship between user and product. In M. Blythe, C. Overbeeke, A. Monk, and P. Wright, editors, *Funology: From Usability to Enjoyment*, 31–42. Dordrecht: Kluwer, 2003.
21. M. Hassenzahl. The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction*, 19(4):319–349, 2004.
22. M. Hassenzahl. User experience (UX): towards an experiential perspective on product quality. In *Proc. IHM '08*, 11–15. ACM, 2008.
23. M. Hassenzahl, A. Beu, and M. Burmester. Engineering joy. *IEEE Software*, 18(1):70–76, 2001.
24. M. Hassenzahl, S. Diefenbach, and A. Göritz. Needs, affect, and interactive products – facets of user experience. *Interacting with Computers*, 22(5):353–362, 2010.
25. M. Hassenzahl and A. Monk. The inference of perceived usability from beauty. *Human-Computer Interaction*, 25(3):235–260, 2010.
26. M. Hassenzahl, A. Platz, M. Burmester, and K. Lehner. Hedonic and ergonomic quality aspects determine a software's appeal. In *Proc. CHI '00*, 201–208. ACM, 2000.
27. M. Hassenzahl and N. Sandweg. From mental effort to perceived usability: transforming experiences into summary assessments. In *Proc. CHI '04*, 1283–1286. ACM, 2004.
28. M. Hassenzahl, M. Schöbel, and T. Trautmann. How motivational orientation influences the evaluation and choice of hedonic and pragmatic interactive products: The role of regulatory focus. *Interacting with Computers*, 20(4-5):473–479, 2008 *.
29. M. Hassenzahl and N. Tractinsky. User experience – a research agenda. *Behaviour & Information Technology*, 25(2):91–97, 2006.
30. M. Hassenzahl and D. Ullrich. To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers*, 19(4):429–437, 2007 *.
31. J. Heikkinen, T. Olsson, and K. Väänänen-Vainio-Mattila. Expectations for user experience in haptic communication with mobile devices. In *Proc. MobileHCI '09*. ACM, 2009 *.
32. M. Hertzum. Images of Usability. *International Journal of Human-Computer Interaction*, 26(6):567–600, 2010.
33. K. Hornbæk. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2):79–102, 2006.
34. K. Hornbæk. Dogmas in the assessment of usability evaluation methods. *Behaviour & Information Technology*, 29(1):97–111, 2010.
35. M. Huang. Web performance scale. *Information & Management*, 42(6):841–852, 2005.
36. ISO. 9241-210: Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. *International Organization for Standardization*, 2008.
37. M. Isomursu, M. Tähti, S. Väinämö, and K. Kuutti. Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *International Journal of Human-Computer Studies*, 65(4):404–418, 2007 *.
38. S. Jackson and H. Marsh. Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of Sport and Exercise Psychology*, 18:17–35, 1996.
39. S. Jacobson and A. Pirinen. Disabled persons as lead users in the domestic environment. In *Proc. DPPI '07*, 158–167. ACM, 2007 *.
40. P. Jordan. *Designing pleasurable products: An introduction to the new human factors*. CRC, 2002.
41. S. Jumisko-Pyykkö, M. Weitzel, and D. Strohmeier. Designing for user experience: what to expect from mobile 3d tv and video? In *Proc. UXTV '08*, 183–192. ACM, 2008 *.
42. E. Karapanos, J. Zimmerman, J. Forlizzi, and J. Martens. User experience over time: an initial framework. In *Proc. CHI '09*, 729–738. ACM, 2009 *.
43. E. Karapanos, J. Zimmerman, J. Forlizzi, and J. Martens. Measuring the Dynamics of Remembered Experience Over Time. *Interacting with Computers*, 22(5):328–335, 2010.
44. A. Koca, M. Funk, E. Karapanos, A. Rozinat, and N. van der Gaarden. Grasping product pragmatics: a case with internet on tv. In *Proc.*

UXTV '08, 193–202. ACM, 2008 *.
45. P. Kourouthanassis, G. Giaglis, and A. Vrechopoulos. Enhancing user experience through pervasive information systems: The case of pervasive retailing. *International Journal of Information Management*, 27(5):319–335, 2007 *.
46. J. Landis and G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.
47. P. J. Lang. Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, and T. A. Williams, editors, *Technology in mental health care delivery systems*, 119–137. Ablex Publishing, Norwood, NJ, 1980.
48. M. Lankes, S. Riegler, A. Weiss, T. Mirlacher, M. Pirker, and M. Tscheligi. Facial expressions as game input with different emotional feedback conditions. In *Proc. ACE '08*, 253–256. ACM, 2008 *.
49. T. Lavie and N. Tractinsky. Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60(3):269–298, 2004.
50. E. Law, V. Roto, M. Hassenzahl, A. Vermeeren, and J. Kort. Understanding, scoping and defining user experience: a survey approach. In *Proc. CHI '09*, 719–728. ACM, 2009.
51. E. Law and P. van Schaik. Modelling user experience – an agenda for research and practice. *Interacting with Computers*, 22(5):313–322, 2010.
52. T. Leong, S. Howard, and F. Vetere. Choice: abidcating or exercising? In *Proc. CHI '08*, 715–724. ACM, 2008 *.
53. A. Lucero, T. Lashina, E. Diederiks, and T. Mattelmäki. How probes inform and influence the design process. In *Proc. DPPI '07*, 377–391. ACM, 2007 *.
54. S. Mahlke. Aesthetic and Symbolic Qualities as Antecedents of Overall Judgements of Interactive Products. *People and Computers XX – Engage*, 57–64, 2007.
55. S. Mahlke and M. Thüring. Studying antecedents of emotional experiences in interactive contexts. In *Proc. CHI '07*, 915–918. ACM, 2007 *.
56. R. Mandryk, K. Inkpen, and T. Calvert. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, 25(2):141–158, 2006 *.
57. N. Mathiasen and S. Bødker. Threats or threads: from usable security to secure experience? In *Proc. NordiCHI '08*, 283–289. ACM, 2008 *.
58. E. McAuley. Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting. *Research Quarterly for Exercise and Sport*, 60(1):48–58, 1989.
59. J. McCarthy and P. Wright. Technology as experience. *Interactions*, 11(5):43, 2004.
60. J. McCarthy, P. Wright, J. Wallace, and A. Dearden. The experience of enchantment in human–computer interaction. *Personal and Ubiquitous Computing*, 10(6):369–378, 2006.
61. D. McKay, S. Cunningham, and K. Thomson. Exploring the user experience through collage. In *Proc. CHINZ '06*, 109–115. ACM, 2006 *.
62. D. Moher, D. Cook, S. Eastwood, I. Olkin, D. Rennie, and D. Stroup. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet*, 354:1896–1900, 1999.
63. M. Montola, T. Nummenmaa, A. Lucero, M. Boberg, and H. Korhonen. Applying game achievement systems to enhance user experience in a photo sharing service. In *Proc. MindTrek '09*, 94–97. ACM, 2009 *.
64. M. Moshagen and M. Thielsch. Facets of visual aesthetics. *Int. Journal of Human-Computer Studies*, 68(10):689–709, 2010.
65. M. Ní Chonchúir and J. McCarthy. The enchanting potential of technology: a dialogical case study of enchantment and the internet. *Personal and Ubiquitous Computing*, 12(5):401–409, 2008 *.
66. D. Norman. *Emotional design*. Citeseer, 2004.
67. H. O'Brien and E. Toms. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6):938–955, 2008 *.
68. H. O'Brien and E. Toms. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50–69, 2010.
69. M. Obrist, A. Meschtscherjakov, and M. Tscheligi. User Experience Evaluation in the Mobile Context. *Mobile TV: Customizing Content and Experience*, 195–204, 2010.
70. Y. Rogers, K. Connelly, L. Tedesco, W. Hazlewood, A. Kurtz, R. Hall, J. Hursey, and T. Toscos. Why it's worth the hassle: the value of in-situ studies when designing Ubicomp. In *Proc. UbiComp '07*, 336–353. Springer-Verlag, 2007 *.
71. P. Ross, C. Overbeeke, S. Wensveen, and C. Hummels. A designerly critique on enchantment. *Personal and Ubiquitous Computing*, 12(5):359–371, 2008 *.
72. K. Scherer. What are emotions? And how can they be measured? *Social Science Information*, 44(4):695, 2005.
73. D. Smith, C. Irby, R. Kimball, and B. Verplank. Designing the Star user interface. *Byte*, 7(4):242–282, 1982.
74. X. Sun and A. May. The role of spatial contextual factors in mobile personalization at large sports events. *Personal and Ubiquitous Computing*, 13(4):293–302, 2009 *.
75. A. Sutcliffe. Assessing the reliability of heuristic evaluation for website attractiveness and usability. In *Proc. HICSS '02*, 1838–1847. IEEE Computer Society, 2002.
76. D. Swallow, M. Blythe, and P. Wright. Grounding experience: relating theory and method to evaluate the user experience of smartphones. In *Proc. EACE '05*, 91–98. ACM, 2005 *.
77. N. Tractinsky, A. Katz, and D. Ikar. What is beautiful is usable. *Interacting with Computers*, 13(2):127–145, 2000.
78. A. Tuch, J. Bargas-Avila, and K. Opwis. Symmetry and aesthetics in website design: It's a man's business. *Computers in Human Behavior*, 26(6):1831–1837, 2010.
79. A. Tuch, J. Bargas-Avila, K. Opwis, and F. Wilhelm. Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human-Computer Studies*, 67(9):703–715, 2009.
80. T. Tullis and W. Albert. *Measuring The User Experience: collecting, analyzing, and presenting usability metrics*. Morgan Kaufmann, 2008.
81. M. Turunen, A. Kallinen, I. Sànchez, et al. Multimodal interaction with speech and physical touch interface in a media center application. In *Proc. ACE '09*, 19–26. ACM, 2009 *.
82. K. Väänänen-Vainio-Mattila and M. Wäljas. Developing an expert evaluation method for user eXperience of cross-platform web services. In *MindTrek '09*, 162–169. ACM, 2009 *.
83. H. Väätäjä, T. Koponen, and V. Roto. Developing practical tools for user experience evaluation: a case from mobile news journalism. In *Proc. ECCE '09*, 1–8. ACM, 2009 *.
84. K. van Boerdonk, R. Tieben, S. Klooster, and E. van den Hoven. Contact through canvas: an entertaining encounter. *Personal and Ubiquitous Computing*, 13(8):551–567, 2009 *.
85. H. Van der Heijden. User acceptance of hedonic information systems. *MIS Quarterly*, 28(4):695–704, 2004.
86. P. van Schaik and J. Ling. Modelling user experience with web sites: Usability, hedonic value, beauty and goodness. *Interacting with Computers*, 20(3):419–432, 2008 *.
87. P. van Schaik and J. Ling. The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies*, 67(1):79–89, 2009 *.
88. P. van Schaik and J. Ling. An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. *Interacting with Computers*, [in press], 2010.
89. R. Wakkary and M. Hatala. ec (h) o: Situated Play in a Tangible and Audio Museum Guide. In *Proc. DIS '06*, page 290. ACM, 2006 *.
90. D. Wilfinger, M. Pirker, R. Bernhaupt, and M. Tscheligi. Evaluating and investigating an iTV interaction concept in the field. In *Proc. EuroITV '09*, 175–178. ACM, 2009 *.
91. P. Wright and M. Blythe. User experience research as an inter-discipline: Towards a UX Manifesto. In *Proc. BHCI '07*, 65–70, 2007.
92. D. Xu, J. Read, G. Sim, B. McManus, and P. Qualter. Children and 'smart' technologies: can children's experiences be interpreted and coded? In *Proc. BCS HCI '09*, 224–231. British Computer Society, 2009 *.

* Denotes a reference among the reviewed studies