Usability, user experience and accessibility: towards an integrative model

Juergen Sauer[1], Andreas Sonderegger[2] & Sven Schmutz[1]

[1]*Department of Psychology, University of Fribourg, Fribourg, Switzerland*

[2]*EPFL+ECAL Lab, École polytechnique fédérale de Lausanne, Renens, Switzerland*

Corresponding author: Juergen Sauer, Department of Psychology, University of Fribourg, Rue de Faucigny 2, 1700 Fribourg, Switzerland; Email: juergen.sauer@unifr.ch

Within the field of ergonomics, the concepts of usability, user experience and accessibility have played an increasingly important role. The present paper examined the meaning of these concepts and their relationship to each other, which included an analysis of the definitions, methods, and typical outcome measures employed. Despite some concerns in the literature about the utility of usability, user experience and accessibility as umbrella terms, we provide arguments for their continued use. The article proposes how the three concepts and their different perspectives can be integrated. We propose the term 'interaction experience' (IX) as a higher-level concept. Due to the multi-facetted nature of umbrella concepts, we suggest using spider charts as a means to report the results of evaluating artefacts with regard to usability, user experience and accessibility.

Keywords: usability; user experience; accessibility; interaction experience; method; ergonomics

Practitioner summary

A better integration of the concepts of usability, user experience and accessibility is expected to provide some benefits to practitioners. We propose employing spider charts for reporting the outcome of artefact evaluations with regard to usability, user experience and accessibility. This may help practitioners interpret the characteristics of a device.

1 Introduction

In the field of ergonomics, the concepts of usability, user experience and accessibility have gained in importance over recent years. They are considered important in ergonomics because of their role in helping design and evaluate interfaces and technical systems (e.g. Kortum 2016). However, these concepts are represented by specific communities comprising researchers and practitioners alike. In particular, there is considerable overlap between the concepts of usability and user experience, which led to a vivid debate in the scientific community about the meaning and utility of these two concepts (e.g. Tractinsky 2018; Bargas-Avila and Hornbæk 2011; Law, Vermeeren, Hassenzahl, and Blythe 2007). Such a debate (though of a less vivid nature) may also characterise the relationship between accessibility and each of the two other concepts (e.g. Petrie and Kheir 2007; Petrie, Hamilton and King 2004; Schmutz, Sonderegger, and Sauer 2016, 2017; Yesilada et al. 2013). While the concepts may be typically used in specific contexts (e.g. usability for everyday products, user experience in software development, and accessibility in the field of housing), we would expect that more significant progress would be made in the field if a more holistic view of the three concepts were adopted by clarifying the mutual benefits in designing user-centred products.

Due the vividness of the debate in some quarters, it may be referred to as the ‘trinity’ of usability, user experience and accessibility. This metaphor does not seem to be that inappropriate given that many beliefs are associated with this issue while empirical facts are playing a minor role.

We will examine in this article the meaning of the concepts of usability, user experience and accessibility and their relationships to each other, which will culminate in a comparative evaluation. This also includes an analysis of the methods, instruments and typical outcome measures employed in research and practice. In the general discussion section of this article we make a proposition of how the three concepts and their slightly different perspectives can be integrated into a single framework. The framework takes into account the predominantly subjective elements of user experience, the mixed conception of usability with both objective and subjective elements, and accessibility with a focus on the consideration of different user groups in designing and testing artefacts. Finally, we propose a new way of assessing the different dimensions of the trinity of usability, user experience and accessibility. In doing this, we will adopt an interactionist perspective (very similar to the one adopted in the ISO definitions of usability and user experience), which takes into account contextual factors (e.g. task, environment, user characteristics, device). This perspective is also reflected in the way we aim to assess the three concepts.

The rationale behind writing this article was to offer some form of integration between the three concepts. A shared view (which is currently lacking) of the three concepts may help advance the field more rapidly because it may avoid fruitless debate over which conceptualisation is most convincing. It is expected to cause less confusion within scientific communities and among practitioners if concepts were used more consistently. We do not claim that the theoretical contribution of the present article will be able to solve the problems at stake but we hope that our article will stimulate some debate of how the concepts can be better integrated.

2 Definition of the three concepts

2.1 Usability

The term usability was coined in the 1980ies as a substitute for the terms ‘user friendliness’ and ‘ease of use’. According to Sarodnick and Brau (2006), the first formal definition of usability was proposed by Shackel (1981) who described the usability of a product as ‘the capability to be used by humans easily and effectively’ (p. 24). Since then a plethora of usability definitions has been proposed, which may be summarised into four different perspectives (Bevan et al. 1991): (a) the *product-oriented perspective* emphasises that usability can be captured as a product-inherent characteristic. (b) According to the *user-oriented perspective*, usability is a function of the user’s mental effort in product usage and their attitude towards the product. (c) Following the *performance-oriented perspective*, usability is described in terms of the interaction of the user with the product. (d) The *context-*

oriented perspective stresses that usability depends on the user group that is studied, the tasks those users are performing, and the environment in which the tasks are completed. It was argued that all those views needed to be taken into account when defining usability (e.g. Bevan et al. 1991). This is the case for the definition of the International Standard Organisation (ISO) which defined usability in its norm ISO FDIS 9241-210 as the 'extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use' (ISO, 2010). Despite a number of differences in the usability definitions, they typically share a number of elements. They generally include objective outcome measures related to performance (e.g. error rate, retention rate, learning rate) but also to subjective outcome measures at the same time (e.g. satisfaction, joy, pleasure). For example, Nielsen (1993) considered an artefact usable if it was easy to learn, efficient, easy to remember, produced few errors and was subjectively pleasing. Other authors defined usability as easy to learn, effective to use and enjoyable from the perspective of the user (Sharp, Rogers and Peerce 2007). Despite these shared elements, usability has been criticised for representing an 'umbrella concept' (Tractinsky 2018). This term refers to a rather broad concept that is loosely employed to describe a set of diverse phenomena (Hirsch and Levin 1999).

When the concept of usability is considered, it is important to make a distinction between the goals and practices of formative and summative testing (Gediga, Hamborg, and Düntsch 1999; Lewis 2014). Formative testing aims to improve a system by identifying the problems in user-product interaction in an iterative product development cycle. Summative testing refers to evaluating the quality of a system with regard to user-product interaction with a view to comparing it against another product or some specified standard. Lewis (2014) even goes as far as suggesting that formative and summative testing represent two distinct conceptions of usability that cannot be integrated into a single concise definition. He refers to 'measurement-based usability' for summative testing and 'diagnostic usability' for formative testing. Summative testing shares considerable elements with practices of experimental psychology whereas formative testing represents a significant departure from it (Lewis 2014). The distinction between formative and summative testing increases the complications associated with the concept of usability further still, adding another layer of complexity.

2.2 User experience

The concept of user experience faces similar problems like usability. User experience is often considered a fuzzy and ill-defined concept (e.g. Law et al. 2007). When examining the many definitions put forward in the research literature, they can be classified into three main types: (a) holistic view of user experience, (b) extension of the usability concept and (c) primary focus on emotion. The 'holistic' approach refers to user experience as all actions, sensations, considerations, feelings and sense making of a person when interacting with a technical device or service (e.g. Dewey 2005; Wright, McCarthy, and Meekison 2003). According to this approach, user experience encompasses a very large range of outcomes of the user-device interaction, including cognitions, attitudes, beliefs, behaviour, behavioural intentions and affect. In the second group of definitions describing user experience as an extension of usability, its scope (i.e. efficiency, effectiveness, satisfaction) is extended by adding affect and emotions. These aspects are not within the scope of the original usability concept. An example for this approach is the Components of User Experience Model (CUE-model, Thüring and Mahlke 2007), which distinguishes between the perception of instrumental qualities (e.g. controllability, effectiveness) and non-instrumental qualities (e.g. visual aesthetics, haptic quality) of a product. A similar model differentiates between hedonic and pragmatic product qualities (Hassenzahl et al. 2008). While pragmatic qualities represent the functional characteristics of the system (i.e. similar to usability), hedonic qualities address the human needs for autonomy, competency, stimulation, relatedness, and popularity. The third approach is similar to the second but concentrates more strongly on the affective outcome of the user-device interaction. For example, Desmet and Hekkert (2007) describe user experience as a set of affective outcomes that are influenced by the experience of aesthetics, emotion and meaning. Similarly, Vermeeren, Kort, Cremers and Fokker (2008) consider user experience as specific emotions (e.g. anger, joy, excitement, unease, and satisfaction), which are influenced by factors such as aesthetics, usability and the personal significance of the product to the user.

The three approaches to defining user experience differ with regard to the scope of outcome measures of user-device interaction. The holistic approach is considered the most encompassing one whereas the third approach (i.e. focus on emotion) represents the most narrow one. It may be a typical feature of scientific theory development that there are different views on the definition of a concept, in the case of user experience, the level of diversity in opinions seems to be particularly high.

With regard to summative and formative evaluation, in principle such a distinction can also be made in user experience. Measuring user experience can focus on qualitative data referring to design aspects that need improving (i.e. formative evaluation) or it can focus on collecting quantitative data about the design with a view to making comparisons to other designs, norms or benchmarks (i.e. summative evaluation). However, in contrast to the usability literature such a distinction has not been explicitly made, to our knowledge, when referring to the term ‘user experience’. Yet it is acknowledged that formative evaluations may have also been carried out when assessing user experience.

Like usability, the concept of user experience found its way into the ISO norms (ISO FDIS 9241-210). It defines the concept as ‘a person's perceptions and responses that result from the use or anticipated use of a product, system or service’ (ISO, 2010). Representing a moderate extension of the scope of the usability definition, it points to the importance of the pre-usage phase (‘anticipated use’) and a broader understanding of user reactions. This raises the question of where the centre of gravity is. Does user experience represent a significant extension of the usability concept or did user experience merely help extend and improve the significance of the usability concept? For example, some authors stress the positive role user experience has played in improving the usability concept (Lewis 2014). Even strong advocates of the concept of user experience acknowledge that ‘… the ideas represented by user experience are important but by no means original’ (Hassenzahl and Tractinsky 2006; p.91). It appears that the understanding of user experience is not only broader but also more diverse (i.e. there is less agreement between scientists) compared to the concept of usability. A similar diagnosis has already been reached some years ago by Bevan (2008), who stated that ‘… the current interpretations of user experience are even more diverse than those of usability’ (p. 17).

2.3 Accessibility

The application area of the term accessibility may be broader than the one of usability and user experience. In addition to the design of websites and consumer products (which are also often examined in usability and user experience), accessibility is also concerned with the design of buildings (e.g. wheel chair access) and transportation (e.g. information for the blind). A strong focus in accessibility research is the design of websites, which has witnessed a considerable amount of work. This focus is due to the importance of barrier-free access to the web for social inclusion of users with impairments. Four main types of impairments are associated with problems in using artefacts: visual, hearing, motor and cognitive impairments. Typical problems in using artefacts include difficulties in reading text with low contrast between text and background, inability to access audio information, and problems in navigating a website with a mouse. A good overview of user characteristics is provided elsewhere (e.g. Alonso-Rios, Vasquez-Garcia, Mosqueira-Rey, and Moret-Bonillo 2010; Ruth-Janneck 2011), including those characteristics in which users may have functional limitations (sensorial, speech, motor and cognitive characteristics). To the groups of users with special needs, one may also include users suffering from mental disorders (e.g. depression, schizophrenia), which may have consequences for their abilities to use artefacts such as the Internet (e.g., Kasckow et al. 2014; Rotondi et al., 2015; Thielsch and Thielsch, 2018).

Accessibility is defined in ISO FDIS 9241-210 as 'the usability of a product, service, environment or facility by people with the widest range of capabilities' (ISO, 2010). This definition extends the concept of usability to a more diverse group of users (e.g. including users with impairments). Similar to user experience but unlike in usability, a distinction between formative and summative testing is not explicitly made in accessibility. Despite the absence of such a distinction, the focus is clearly on methodological approaches similar to summative testing, though the methods are somewhat different from the classic usability test (see also 3.3 for a more detailed treatment).

Related to the concept of accessibility are the concepts 'design for all' (e.g. Benedixen 2015), 'inclusive design' (e.g. Petrie and Edwards 2006) and 'universal design' (e.g. Iwarsson and Stahl 2003). Like accessibility, the three concepts emphasise the idea of designing artefacts that are usable for users with a very wide range of capabilities (which naturally includes users with impairments). Unlike accessibility, these concepts do not focus on designing for special user groups but rather emphasise the need to consider people with and without impairments. However, designing for the widest range of users is a difficult endeavour because some user groups may have contradicting needs. For instance, whereas people with cognitive impairments benefit from presenting pictorials on websites, users who are blind may have difficulties because typical screen reading software cannot interpret pictorials (Thatcher et al. 2006). Therefore, designing for different user groups often requires a compromise. Such a compromise may be in conflict with the idea of designing for special groups of users (see Petrie and Bevan 2009). Therefore, it is important to determine whether designing for a specific user group will only provide benefits to the target group or whether other users (e.g. nondisabled users) would benefit, too (in the sense of 'universal design'). There seems to be no unequivocal answer to this question. Recent research evidence suggests that the implementation of web accessibility guidelines may provide benefits for nondisabled users, too (Schmutz et al. 2016, 2017, 2018), but in specific areas (e.g. using easy language) the benefits for nondisabled users are limited (Schmutz et al. 2019).

2.4 Comparative analysis and conclusion

The analysis of the three concepts has indicated that usability, user experience and accessibility share some conceptual similarities but also differ with regard to particular aspects. This analysis was based on a literature review that does not fulfil the criteria for a systematic review but it reflects our understanding of how the field was conceptualised. In a first step, we compare usability to user experience. In a second step, accessibility will be compared to usability and user experience, with a view to integrating the three concepts into a common framework in a later step.

First, the comparison of usability and user experience showed that the similarities and differences vary considerably as a function of the type of definition chosen for user experience. The implications of these differences in definition are summarised in figure 1, reflecting our understanding of the conceptualisations found in the research literature. If user experience is defined according to the 'holistic' and thus broadest approach, it encompasses all subjective, emotional and behavioural consequences of the interaction of a user with an artefact. Since such a definition includes also satisfaction and user behaviour (i.e. performance in this context), user experience represents an extension of the usability concept incorporating all components of the usability concept (see figure 1a). If the second approach is adopted, the definition of user experience encompasses all subjective components of the user-artefact interaction but does not include the 'performance/behaviour' component. User experience and usability would therefore describe two separate concepts, which share the 'satisfaction/subjective evaluation'-component (see figure 1b). Following the third, and most narrow definition of user experience as merely the affective outcome of the user-artefact interaction, user experience and usability represent two independent concepts sharing no components with each other (see figure 1c). We acknowledge that the concepts of affect and satisfaction are not mutually exclusive (see also Lindgaard and Dudek, 2003). The term satisfaction as an attitude may be considered as consisting of a cognitive and an affective component (e.g. Kaplan, Warren, Barsky, & Thoresen, 2009). Despite some degree of overlap between the two concepts, we believe that they are sufficiently different to justify their representing separate entities in the model.

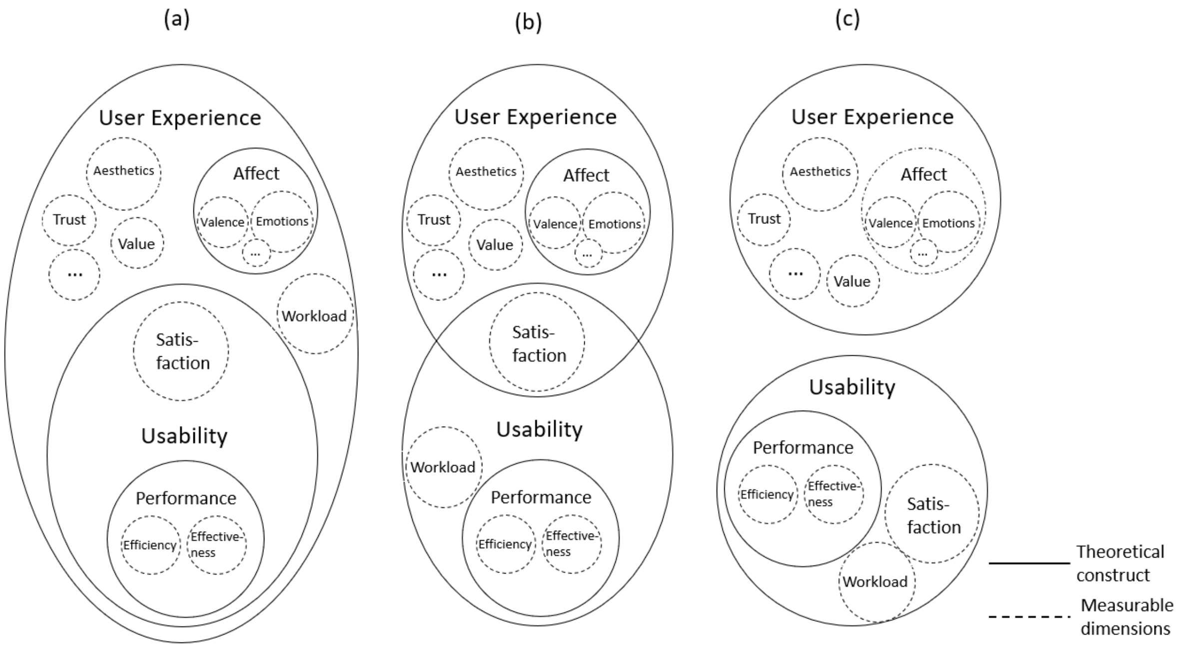

Figure 1. Different views on the relationship between usability and user experience: (a) broad-scope definition of user experience, (b) medium-scope definition of user experience, (c) narrow-scope definition of user experience

Second, comparing accessibility to usability and user experience shows that accessibility may be more closely related to usability than to user experience (as the ISO definition clearly demonstrates). According to the definition of the ISO, accessibility considers the usability of an artefact by focusing on a very wide range of users with different requirements. Comparing this definition of accessibility to the ISO-concept of usability, in our understanding it implies that all users (with and without special needs) should be able to interact with an artefact effectively, efficiently and satisfactorily. However, this definition does not include the affective part that is generally covered by the definition of user experience, thus implicitly implying that fun, enjoyment and other affective outcomes of user interaction are not important for users with impairments. This implies that accessibility can benefit from user experience by establishing a closer link between the two concepts. This would allow designers to develop usable artefacts that are, at the same time, fun and exciting to use. Such a goal should not be limited to the large group of non-disabled users but should also apply to users with functional impairments. Therefore, affective outcomes of the user-artefact interaction need to be integrated into the accessibility concept, too.

The analysis also revealed that the distinction between formative and summative testing in usability research and practice is hardly made explicitly in the research literature concerned with the fields of accessibility and user experience, though in practice formative testing may be used in product development in both fields (yet to a much smaller extent in the field of accessibility). The fact that the theoretical distinction between formative and summative testing is largely absent in these literatures may not be to the benefit of the two fields. In particular, in accessibility the focus is clearly on summative testing. Therefore, some gains could be made in research and practice if formative testing became more common. Since formative testing focuses on an iterative design process with the aim of identifying issues that could be improved, such testing might be advantageous. For instance, practitioners typically argue that implementing accessibility involves too much effort and too high costs (e.g. Thatcher et al. 2006). However, we believe that it would represent an excellent investment to consider accessibility already early in the development process rather than only towards the end.

3 Methodological approaches used when applying the three concepts

3.1 Usability evaluation

The literature provides a plethora of methods that aim to capture the usability of artefacts. These include user testing, observation, questionnaires, interviews, focus groups, heuristic evaluation, cognitive walkthrough and data logging (Nielsen 1993). The research literature provides several articles, in which overviews and classification of the usability methods may be found. This includes generic applications (Hartson, Andre, and Williges 2003) but also specific applications such as virtual environments (Bowman, Gabbard, and Hix 2002) and software (Fitzpatrick 1998). Most of these methods have their origins in usability evaluation, in which they are still predominantly used. In usability evaluation, both expert-based and user-based methods are generally considered valuable. Expert-based methods are considered cost-effective and economical, being mainly used in formative evaluation. However, expert-based methods have been criticised for not considering the user and their needs and limitations in a sufficient way. In addition, the subjective component of usability is difficult to assess through expert methods. Therefore, user-based methods are an essential part for usability evaluation. The principal method to capture the usability of artefacts represents the user test (Lewis 2006), which is often referred to as the usability test. In a usability test, typical users interact with an artefact while being watched by one or several observers (Kuniavsky 2003). Such observations provide quantitative and objective data of user behaviour (e.g. efficiency, effectiveness, errors and usage problems). In addition, participants are often requested to verbalise their thoughts while interacting with the artefact in order to provide additional information to the evaluator (i.e. thinking-aloud technique). In addition to these subjective data, a usability test is usually followed by a set of questionnaires to capture concepts such as user satisfaction and perceived usability. A number of tools have been developed over the years to measure perceived usability and related concepts by means of questionnaires. This includes PSSUQ (Post-Study System Usability Questionnaire; Lewis 1995), SUS (System Usability Scale; Brooke 1996), IsoMetrics (Gediga et al. 1999), QUIS (Questionnaire for User Interface Satisfaction; Chin et al. 1988), SUMI (Software Usability Measurement Inventory; Kirakowski, 1996), and WAMMI (Web Analysis and Measurement Inventory; Kirakowski and Cierlik 1998). These instruments represent well-established measurement tools, with some of them also enjoying good psychometric properties. The vast majority of usability tests are carried out as one-off testing sessions (e.g. Sonderegger, Zbinden, Uebelbacher, and Sauer 2012), which is problematic. Hornbæk (2006) raises the important issue of time-dependent changes in usability as users gain increasing experience with the system. This is of obvious importance for practice because many products are used for over a long period. Despite the practical importance of long-term effects, there is a lack of long-term studies in this area. An exception is the study of Kujala et al. (2011), which proposed a qualitative method to evaluate long-term user experience, emphasising the important role of long-term process for device acceptance.

In conclusion, it emerged that a very broad range of methods are used to assess usability. While expert-based methods provide mainly one sort of data (e.g. usability issues, suggestions for improvement), user-based methods often combine several dimensions of the usability concept (e.g. effectiveness, satisfaction). This combined approach results in a broader range of data being collected (e.g. task completion time, error rate, task completion rate, satisfaction and attitudes).

3.2 User experience evaluation

In contrast to the large spectrum of tools employed in usability evaluation, methods and instruments for measuring user experience are smaller in number and less diverse. Expert-based methods such as cognitive walkthrough, heuristic analysis and checklists are often considered inappropriate because user experience is believed to be a very individual reaction of a user to an interface (Law et al. 2007). This is because it is very difficult for experts to predict such an individual reaction by cognitively walking through the interface, asking themselves the question whether a potential user might experience fun, frustration or other emotion when interacting with the device. In addition, since user experience depends on the very individual needs, expectations and limitations of users, it seems to be rather difficult to define and use universally applicable checklists and guidelines. Therefore, user-based methods are predominantly applied in user experience evaluation. The type of data assessed varies as a function of the definition chosen for user experience (see section 2.2). Since it is impossible to do justice to the complexity of the concept of user experience within the realm of the present article, the reader is also referred to other sources (e.g. Bargas-Avila and Hornbæk 2011).

Considering user experience as an affective outcome of the user-artefact interaction (see figure 1c), the concept can be measured by assessing the emotional reactions of a user. In the measurement of user experience, emotions are usually captured by means of verbal or non-verbal questionnaires such as the self-assessment manikin (SAM; Lang 1980), the Geneva Emotion Wheel (Sacharin, Schlegel, and Scherer 2012) or the product emotion tool (PREMO; Desmet, Hekkert, and Jacobs 2000). In addition to measuring emotion as a specific aspect of user experience, several tools have been developed in the form of self-report questionnaires to capture user experience as a comprehensive concept. This approach of measuring the umbrella concept with one tool is similar to the approach of measuring usability with self-report questionnaires, though user experience questionnaires are somewhat smaller in number. Prominent examples of such instruments are UEQ (User Experience Questionnaire; Laugwitz et al. 2008) and meCUE (Modular Evaluation of Key Components of User Experience; Minge and Riedel 2013). These instruments are considered to have good psychometric qualities. While user experience largely relies on subjective and quantitative measurements, emotion (as one aspect of user experience) can also be measured more objectively by using physiological parameters (e.g. Mandryk, Inkpen, and Calvert 2006), body posture (Tan et al. 2013), and facial and vocal cues (e.g. Fasel and Luettin 2003). In affective sciences, considerable effort has been expended developing new methods that assess affective states automatically. While such objective methods may be useful for complementing subjective questionnaire data, it is less clear whether they would be more suitable than the subjective methods to capture user's affective state. For practitioners in particular, these objective methods come at several disadvantages, such as intrusiveness of measurement devices, invasion of privacy, sophisticated and time-consuming data analysis, and difficult interpretation of the results.

With user experience being considered as an extension of the usability concept (see figure 1a), the concept tends to be assessed by means of questionnaires. A number of questionnaires have been published that aim to assess the two components of the user experience (usable / pragmatic vs. affective / hedonic), such as the AttrakDiff (Hassenzahl et al. 2008) and the User Experience Questionnaire (UEQ; Laugwitz, Schrepp, and Held 2006). The AttrakDiff differentiates between pragmatic quality (i.e. usability) and identification, stimulation and attractiveness (i.e. affective/hedonic qualities). Similarly, the UEQ suggests six dimensions, with efficiency, perspicuity, and dependability being strongly linked to the usability component while stimulation, novelty, and attractiveness are considered part of the affective/hedonic component. One may even argue that in early usability scales affective components were already assessed by certain items (e.g. SUMI: Kirjakowsky, 1996; QUIS: Chin et al. 1988). These two-component models have been enlarged more recently to four components, which include product perception (e.g. perceived usability, aesthetics), user emotions (e.g. positive and negative ones), consequences of usage (e.g. intention to use), and an overall judgement of the product (Thüring and Mahlke 2007; Minge et al. 2016). This resulted in the development of a questionnaire (called *meCUE*; Minge et al. 2016), which aims to capture the broader, four-component concept of user experience. A similar idea to the four-component concept of user experience is also reflected in the broad approach of measuring user experience, depicted in figure 1a. This approach defines the concept as a user's comprehensive experience in the context of their interaction with a device or service. This may include emotions, trust, workload, intention to buy and intention to recommend the device. It is clearly more difficult to measure user experience following such a broad-scope definition than when using a narrow-scope definition.

3.3 Accessibility evaluation

In contrast to usability and user experience, user testing is not typically carried out in the accessibility. Accessibility focuses more on expert-based methods rather than methods involving the user. Three methods are predominantly used. First, *checklists* provide specific recommendations about how to support users with impairments (e.g. Web Content Accessibility Guidelines 2.0; Caldwell et al. 2008). Second, *cognitive (barrier) walkthroughs* aim to identify barriers by using severity and by naming performance attributes that may be affected (e.g. Brajnik 2008; Yesilada, Brajnik, and Harper 2009). Third, *automatic checking* measures quantifiable parameters (e.g. colour contrast between text and background to judge perceptibility) by means of algorithms (www.achecker.ca/checker; www.wave.webaim.org). In contrast to usability and user experience, for accessibility there are no tools available in the form of self-report questionnaires.

There are specific reasons for the low prevalence of user testing in accessibility evaluation. Usability engineers have often argued that testing with people with impairments requires too much time and too many financial resources (Farrely 2011; Petrie and Kheir 2007; Rutter et al. 2007). Furthermore, it may be difficulty to gain access to users with functional limitations. The reservations on the side of accessibility practitioners to carry out user testing may be partly due to uncertainties about being able to cater for the special needs of users with impairments during the testing procedure (e.g. users with cognitive impairments). Our own experience in testing users with cognitive impairments suggests that testing procedures need to be strongly adapted to meet the needs of this special user group (e.g. users with cognitive impairments require rather short testing times and find it difficult to report their subjective evaluation of the artefact). The questionnaires available may need to be adapted for different accessibility groups. This may include users with cognitive impairments (e.g. easy language needs to be used), users with motor impairments (e.g. response format needs to be adapted) and users with visual impairments (e.g. paper format may be inappropriate). The problem of gaining access may be partly overcome by contacting charities and other organisations that support the interests of different groups of users with functional limitations. Despite the difficulties associated with testing users with special needs, several authors have emphasised the importance of user testing in the field of accessibility (Henry 2006; Thatcher et al. 2006). There are even a few examples where it was successfully employed (e.g. Waller, Bradley, Hosking, and Clarkson 2015).

Overall, we do not think that the arguments against user testing are convincing. They are reminiscent of similar arguments used by traditional managers and product developers in the 1980ies when usability testing emerged (e.g. Shneiderman, 1998). At that time, many of them considered testing the device in a lab with prospective users to be an endeavour not worth the effort because they were very confident that they would know best what a usable artefact should look like. Nowadays we do not hear much about this argument in the field of usability evaluation because the methodological approaches have improved. This clearly raises the question of whether the arsenal of methods should not be enlarged in accessibility evaluation, too. We believe that the field of accessibility would benefit if a similar development took place, at the end of which it will be evident that in addition to the guideline-centred methods, user-centred ones are also needed to improve the accessibility of products. This is also related to the theoretical deliberations about the appropriate scope of accessibility (see section 2.4). We believe that accessibility needs to include the notion of designing for positive experiences and emotions, going beyond a mere functional design (e.g. focussing on efficiency and effectiveness). In other words, accessible design should also offer to users some fun, enjoyment and positive emotions when interacting with the device. The measurement of such affective experience cannot be achieved by using expert-based methods (see section 3.2). Therefore, user-based evaluation methods need to become a prominent as part of evaluation procedures in accessibility.

3.4 Comparative analysis and conclusion

The comparative analysis revealed that the three concepts have been measured by a large range of methods, with considerable differences emerging between the methods used when applying the three concepts. The largest range of methods was used in usability evaluation, which employed both user-based and expert-based methods alike (though the focus may be slightly more on user-based methods). User experience was characterised by a strong tendency to employ user-based methods, focusing particularly on subjective data. In contrast, accessibility focuses very much on expert-based method.

Table 1 provides an overview of the methods used when applying the three concepts. Based on the preceding literature review, a coarse assessment is provided by the authors of how often the different evaluation methods are typically employed. This assessment is based on the literature review carried out for writing this article but also the authors' general expertise in this area. It makes use of broad categories rather than aim to quantify the rating of each criterion based on a quantitative analysis. Table 1 provides a description of the type of data predominantly used for each method. It may be helpful to classify the methods along the following three dimensions: data source (expert vs. user), quantitative/qualitative data and subjective/objective data. The table highlights four important points. First, it shows that the range of methods used is most varied in the usability evaluation. In contrast, accessibility shows the least varied use of different methods, with intermediate levels being observed for user experience. Second, it shows that user-based methods are more frequently used in usability and user experience than in accessibility, where user-oriented approaches are least frequently employed. Third, accessibility uses objective methods more frequently than subjective ones. The reverse picture emerges for user experience. Usability evaluation is characterised by the most balanced collection of both types of data. Fourth, the table shows that in the accessibility, quantitative data are more often collected than qualitative data. This is in contrast to both usability and user experience, which show that the use of both types of data is relatively balanced. Overall, the comparative rating of the method usage reveals that the usability is characterised by a more balanced use of the different methods and different types of data than the two others.

Table 1: Overview of methods employed in the three domains (***frequently employed, **occasionally employed, *hardly employed[1])

Method	Type of data	Usability	User experience	Accessibility
User testing	UB, QL, QNT, OBJ	***	***	*
Checklist	EB, QNT, OBJ, SUBJ	**	*	***
Heuristic evaluation	EB, QL, QNT	**	*	**
Questionnaire	UB, QNT, SUBJ	***	**	*
User reports and complaints	UB, QL, SUBJ	***	**	**
Thinking-aloud technique	UB, QL, SUBJ	***	**	*
Cognitive walkthrough	EB, QL, SUBJ	**	*	*
Automatic checking	EB, QNT, OBJ	*	*	***
User observation	UB, QL, QNT, OBJ	***	**	*
Psychophysiology	UB, QNT, OBJ	**	**	*

UB: user-based, EB: expert-based, QL: qualitative, QNT: quantitative, OBJ: objective, SUBJ: subjective.
[1]Assessment is based on authors' judgement.

The theoretical analysis of the three concepts showed that usability, user experience and accessibility are of a multidimensional nature. This makes it difficult to measure them. It raises the question of whether we should aim to compute a single aggregated score for each of the three concepts, or even for all three concepts together. Such an aggregated score for each concept would have the advantage of offering an easy-to-understand estimate of the qualities of the artefact.

For the usability construct, a meta-analysis concluded that there are rather low correlations between subjective and objective measures (Hornbæk and Law, 2007). In contrast, other work provided some empirical evidence in support of a 'g-factor' (i.e. an underlying principal component representing general usability), suggesting that a combination of the different dimensions of usability to one usability measure might be acceptable (Sauro and Lewis, 2009).

For user experience, there is some empirical evidence that may be counted against a 'g-factor', with correlations between sub-dimensions of user experience being of rather small magnitude (e.g. Sonderegger et al., 2019). The answer to question about the 'g-factor' may also depend on the definition of user experience being adopted. If the narrow definition of user experience is chosen (see figure 1c), it appears more appropriate to aggregate the sub-dimensions into a single score than if the broad and holistic conceptualisation of user experience was chosen (see figure 1a). Overall, there is little empirical data available to support one view or the other.

To take our deliberations a little further, we also need to raise the question whether we should aim for a single score or multidimensional scores when we attempt to assess an artefact with regard to two of the concepts (e.g. usability and user experience) or even all three of them. There are no data available yet to support or discount the idea of aggregating the assessment of the three concepts of usability, user experience and accessibility into a single score or into multiple scores.

To reconcile the two positions, we would suggest that an aggregated single score could be provided alongside a representation of the scores on each dimension. This would also be in line with the position of Sauro and Lewis (2012), who argue that no disadvantage is associated with combining scores because it merely provides an additional view without losing any of the original constituting data permanently.

If multiple scores of the concepts of usability, user experience and accessibility are to be assessed, the measurement of these dimensions should make use of existing questionnaires wherever possible (cf. Sonderegger et al. 2019). The reporting could be done in the form of a spider chart (see figure 2). The spider chart allows, at a glance, a good understanding of the strength and weaknesses of the device as a function of the different dimensions. It would provide similar advantages as configural displays used in human-machine interaction (e.g. Bennett, Toms, and Woods 1993). Based on the shape of the graph, one could easily see what the strength and weaknesses of the device are. For example, figure 2a shows that the two devices differ considerably in user experience, usability and accessibility, though the overall score is the same for both. The regular usage of standardised spider charts could help practitioners interpret the characteristics of a device at a glance. The particular shape of the chart would allow domain experts to detect immediately whether the device complies with their expectations or does not. Such a multidimensional understanding of the concepts may also provide important benefits to practitioners because it may reveal useful information when evaluating a product or deciding about specific characteristics of a product in the design process (see figure 2b). While device 1 scores high on most aspects except for efficiency, device 2 shows the opposite pattern by scoring very high on efficiency but rather low on the other aspects. If we had considered the aggregated score of usability, device 1 would have been the clear winner because the aggregated score (i.e. total surface inside the polygon chart) is larger than for device 2. However, in addition to the multiple scores, the aggregated score can also be taken into account for the evaluation of the product (e.g. by examining the total inside the polygon chart or by providing a single score as a number).

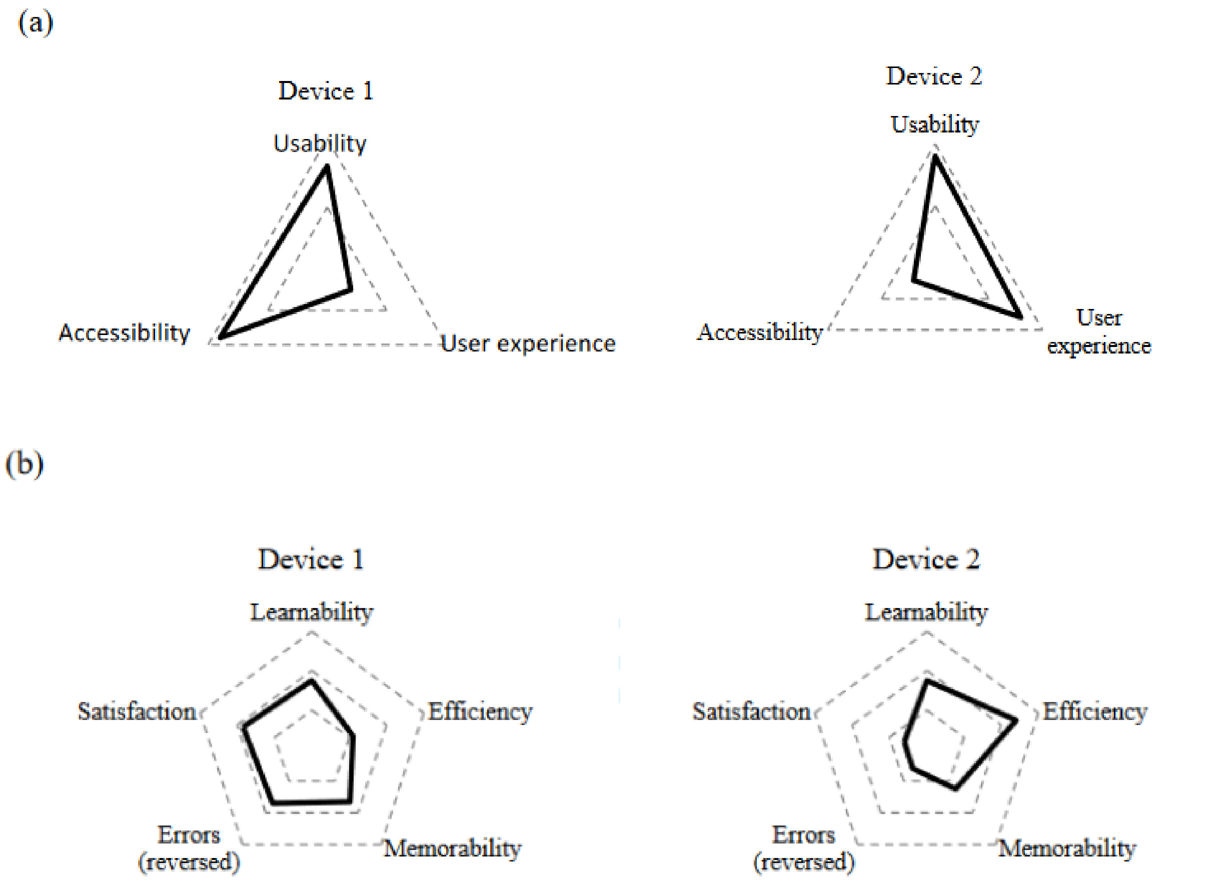

Figure 2: Four examples of fictitiousspider charts (a) in the form of a triangle depicting the three elements of ‘interaction experience’ (i.e. user experience, usability and accessibility), and (b) in theform of a pentagon depicting Nielsen’s (1993)five elements of usability.

4 General discussion and conclusion

Having reviewed the research literature and carried out the comparative analysis of the three concepts, we agree with some of the criticism levelled at the concepts of usability, user experience and accessibility but we also believe that a more balanced view needs to be adopted by more strongly acknowledging the benefits of using these concepts. Based on the deliberations in the present article, we would like to make a new proposition of how the relationship between the terms usability, user experience and accessibility may be defined. The basic structure of the relationship is presented in figure 3. The figure acknowledges the importance of the many facets of user experience and usability but also emphasises the important contribution of accessibility.

We propose ‘interaction experience’ (IX) as a new higher-level concept, which encompasses both user experience and usability. We also propose that this new term IX incorporates the important ideas conveyed by accessibility. We believe that there is a need to suggest a new concept for the following reason. Given the vivid debate about the definitions of user experience and usability and how they relate to each other (with the different views being summarised in figures 1a-c), it is difficult to see that the scientific community would agree on a conceptual framework, in which either user experience or usability would emerge as the higher-level term encompassing the other one.

We are aware of the problem of introducing another concept into a field that is characterised by some inconsistencies in the use of scientific terms. However, in this case we firmly believe that such a new concept is necessary for two reasons. First, it could alleviate the problem of the excessive use of the fuzzy and ill-defined UX concept (Law et al., 2007). Second, it may reconcile the opposing views surrounding the definitions of usability and UX (‘Usability is a part of UX!’ versus ‘UX is a modest extension of usability!’). Although the term IX has already been used in the research literature on a very few occasions (e.g. van Schaik and Ling, 2011), it has not been employed consistently with the same specific meaning attached to it. Therefore, we think that IX as a new concept may contribute to turning the continuing debate about the role of usability, user experience and accessibility into a positive direction.

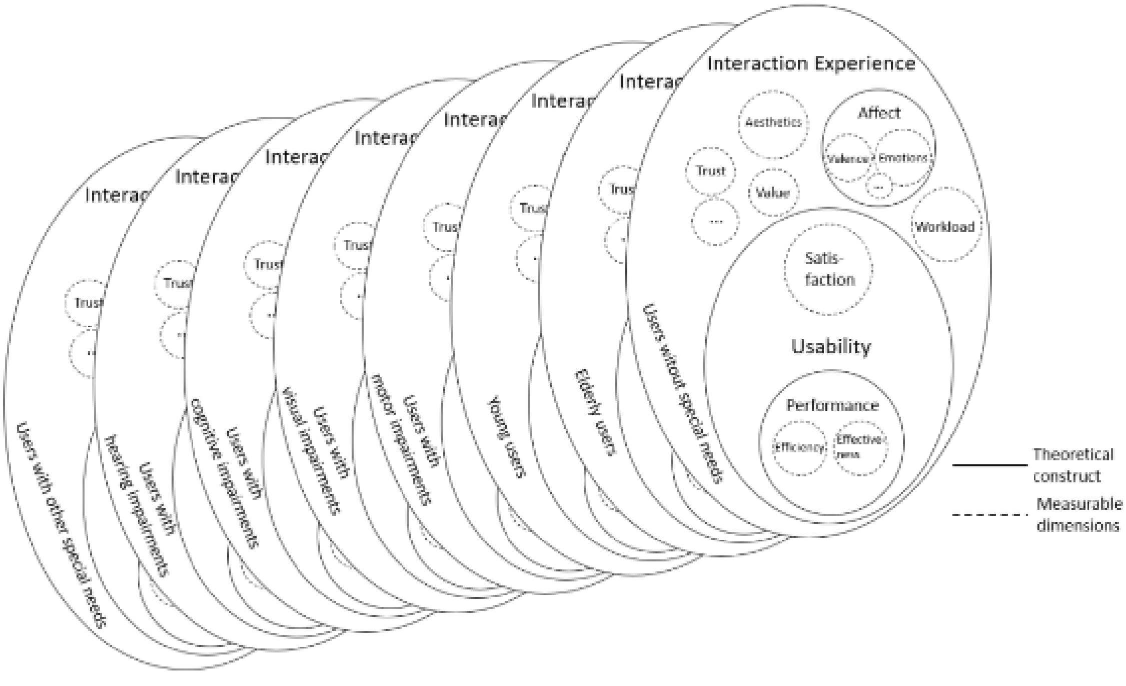

Figure 3. Relationship of the concepts of usability, user experience and accessibility to the higher-level concept 'interaction experience' (IX)

The drawing in figure 3 emphasises a holistic view of IX as the concept encompassing all facets of the experience a user is gaining when interacting with an artefact. IX as a general term cannot be measured. It is considered an umbrella term. However, its constituting elements (accessibility, efficiency, user satisfaction, emotion, etc.) can be measured by making use of established instruments and methods (e.g. questionnaire or usability test). There are two reasons for assuming that the model may provide benefits to the scientific community. First, in future research a clear distinction between different concepts would be advantageous because it allows a more precise examination of issues in user-artefact interaction. Second, for practitioners such a clear distinction would be helpful because it draws their attention to the wide range of aspects that need to be addressed, including satisfaction, emotion and aesthetics. Figure 3 also emphasises the importance of considering different user groups that need to be taken into account during artefact development. Identifying all these user groups and their respective needs is a major contribution of accessibility. Recent empirical work has emphasised the similarities between accessibility and the other two concepts (Schmutz et al. 2016, 2017, 2018, 2019). We believe that integrating the three concepts into a common framework provides the advantage of being able to make use of the strengths of each concept (e.g. by highlighting certain issues). We acknowledge that the utility of the model proposed needs to be examined by researchers and practitioners alike in future work.

The conceptual framework proposed enjoys the additional advantage of matching the semantic connotations of the terms ‘user experience’ and ‘usability’. User experience connotes a strong subjective perspective, which may have neglected the measurement of objective parameters such as user performance. This is due to the term ‘experience’, which has a clear subjective connotation. This is in contrast to the term ‘usability’, which has a more neutral connotation and is associated with subjective and objective measures alike. The semantic connotation of IX makes it possible that both usability and user experience are embodied in the term. The first word in the compound IX reflects the strong emphasis of the ‘usability’ concept on the quality of the interaction between the user and the artefact while the second word takes into account the important experiential aspect of the term ‘user experience’.

In summary, we believe that the present article makes a useful contribution to the on-going debate about the role of the concepts of usability, user experience and accessibility. This article suggested ways towards a better integration of the concepts. This will provide benefits by broadening the range of methods used to measure each concept (e.g. accessibility could benefit from usability by making more use of user-based methods such as empirical testing with prospective users). An analysis of the different methodological approaches associated with each concept has been summarised in table 1, suggesting ways of transferring methods to other fields. Furthermore, the article provides an overview of the different definitions of the three concepts within the scientific community, highlighting the differences in opinions in the field. We than made a proposition of how the three concepts and their different perspectives can be integrated, suggesting the term IX as a higher-level concept. To overcome the problems of measuring umbrella concepts, we suggest using spider charts as a means to report the results of evaluating artefacts with regard to usability, user experience and accessibility.

References

Alonso-Ríos, D., A. Vázquez-García, E. Mosqueira-Rey, and V. Moret-Bonillo. 2010. A context-of-use taxonomy for usability studies. *International Journal of Human-Computer Interaction 26*(10): 941-970.

Bargas-Avila, J. A., and K. Hornbæk. 2011. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 2689-2698. ACM.

Bennett, K. B., M. L. Toms, and D. D. Woods. 1993. Emergent features and graphical elements: Designing more effective configural displays. *Human Factors* 35(1): 71-97.

Bendixen, K., and M. Benktzon. 2015. Design for All in Scandinavia – A strong concept. *Applied Ergonomics* 46: 248-257.

Bevan, N. 2008. Classifying and selecting UX and usability measures. *International Workshop on Meaningful Measures: Valid Useful User Experience Measurement*. 13-18.

Bevan, N. K., J. Kirakowski, and J. Maissel, 1991. What is usability? *Proceedings of the 4th International Conference on HCI,* Stuttgart.

Bowman, D. A., Gabbard, J. L., and Hix, D., 2002. A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence: Teleoperators & Virtual Environments,* 11(4), 404-424.

Brajnik, G. 2008. Beyond conformance: the role of accessibility evaluation methods. *Web Information Systems Engineering–WISE 2008 Workshops*: 63-80. Berlin/Heidelberg: Springer.

Brooke, J. 1996. SUS-A quick and dirty usability scale. *Usability Evaluation in Industry* 189(194): 4-7.

Caldwell, B., M. Cooper, L. G. Reid, and G. Vanderheiden. 2008. *Web content accessibility guidelines 2.0, World Wide Web Consortium (W3C) recommendation*. Retrieved from http://www .w3.org/TR/2008/REC-WCAG20-20081211/

Chin, J. P., Diehl, V. A., and Norman, K. L. 1988. Development of an instrument measuring user satisfaction of the human-computer interface. In Proceedings of the SIGCHI conference on Human factors in computing systems, 213-218.

Davis, F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 319-340.

Desmet, P., and P. Hekkert. 2007. Framework of product experience. *International Journal of Design* 1(1).

Desmet, P. M., P. Hekkert, and J. J. Jacobs. 2000. When a car makes you smile: Development and application of an instrument to measure product emotions. *ACR North American Advances*.

Dewey Jr, R. G. 2005. *U.S. Patent No. 6,938,634*. Washington, DC: U.S. Patent and Trademark Office.

Farrelly, G. 2011. Practitioner barriers to diffusion and implementation of Web accessibility. *Technology and Disability* 23: 223–232. doi: 10.3233/TAD-2011-0329

Fasel, B., and J. Luettin. 2003. Automatic facial expression analysis: a survey. *Pattern recognition* 36(1): 259-275.

Fitzpatrick, R. 1998. Strategies for evaluating software usability. http://www.comp.dit.ie/rfitzpatrick/papers/chi99%20strategies.pdf

Gediga, G., K. C. Hamborg, and I. Düntsch. 1999. The IsoMetrics usability inventory: an operationalization of ISO 9241-10 supporting summative and formative evaluation of software systems. *Behaviour & Information Technology* 18(3): 151-164.

Hart, S. G., and L. E. Staveland. 1988. Development of NASA- TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology* 52: 139–183. doi: 10.1016/ S0166-4115(08)62386-9

Hartson, H. R., Andre, T. S., & Williges, R. C. 2003. Criteria for evaluating usability evaluation methods. *International Journal of Human-Computer Interaction*, *15*(1), 145-181.

Hassenzahl, M., M. Burmester, and F. Koller. 2008. "Der User Experience (UX) auf der Spur: Zum Einsatz von www.attrak.diff." In *Usability Professionals*, edited by H. Brau, S. Diefenbach, M. Hassenzahl, F. Koller, M. Peissner, and K. Rose, 78-82. Germany: Frauenhofer IRB Verlag.

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

Henry, S. L. 2006. Understanding Web accessibility. In *Web accessibility: Web standards and regulatory compliance*, edited by J. Thatcher, R. M. Burks, C. Heilmann, S. L. Henry, A. Kirkpatrick, P. H. Lauke, B. Lawson, B. Regan, R. Rutter, M. Urban, and C. D. Waddel, 1-51. New York, NY: Apress.

Hirsch, P. M., and D. Z. Levin. 1999. Umbrella advocates versus validity police: A life-cycle model. *Organization Science* 10(2): 199-212.

Hornbæk, K. 2006. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies* 64(2): 79-102.

Hornbæk, K., & Law, E. L.-C. 2007. Meta-analysis of Correlations Among Usability Measures. Proceedings of the SIGCHI Conference on Human Factors in Computing

Systems, 617–626. https://doi.org/10.1145/1240624.1240722

International Organization for Standardization. (2010). Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems (Standard No. 9241-210). Retrieved from http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/20/52075.html

Iwarsson, S., and A. Ståhl. 2003. Accessibility, usability and universal design—positioning and definition of concepts describing person-environment relationships. *Disability and rehabilitation* 25(2): 57-66.

Kaplan, S. A., Warren, C. R., Barsky, A. P., and Thoresen, C. J. 2009. A note on the relationship between affect (ivity) and differing conceptualizations of job satisfaction: Some unexpected meta-analytic findings. *European Journal of Work and Organizational Psychology* 18(1), 29-54.

Kasckow J., Zickmund S., Rotondi A., Welch A., Gurklis J., Chinman M., Fox L., Haas G.L. 2014. Optimizing scripted dialogues for an e-health intervention for suicidal veterans with major depression. *Community Mental Health Journal* 51(5):509–512.

Kirakowski, J. 1996. The Software Usability Measurement Inventory: Background and usage. In *Usability Evaluation in Industry*, P. Jordan, B. Thomas, and B. Weerdmeester (Eds.), pp. 169-178. London, UK: Taylor & Francis.

Kirakowski, J., and Cierlik, B. 1998. Measuring the usability of web sites. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 42 (4): 424-428. Sage CA: Los Angeles: SAGE Publications.

Kortum, P. 2016. *Usability assessment: how to measure the usability of products, services, and systems.* USA: Human Factors and Ergonomics Society.

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

Kuniavsky, M. 2003. *Observing the user experience*. San Francisco: Morgan Kaufman.

Lang, P. 1980. “Behavioral treatment and bio-behavioral assessment: computer applications.” In *Technology in mental health care delivery systems*, edited by J. Sidowski, J. Johnson, and T. Williams, 119-137. Norwood, NJ: Ablex.

Laugwitz, B., M. Schrepp, and T. Held. 2006. Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. *Mensch & Computer*: 125-134.

Laugwitz, B., Held, T., & Schrepp, M. (2008, November). Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group* (pp. 63-76). Springer, Berlin, Heidelberg.

Law, E. L.C., A. P. O. S. Vermeeren, M. Hassenzahl, and M. Blythe. 2007. "Towards a UX Manifesto." In *HCI… But Not As We Know It*, Proceedings of the 21st British HCI Group Annual Conference on People and Computers, 205-206. Swindon, UK; BCS Learning & Development Ltd.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57-78.

Lewis, J.R. 2006. Usability testing. In: Salvendy, G. (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 1275–1316). New York: John Wiley.

Lewis, J. R. 2014. Usability: lessons learned… and yet to be learned. *International Journal of Human-Computer Interaction* 30(9): 663-684.

Lindgaard, G., and C. Dudek. 2003. What is this evasive beast we call user satisfaction? *Interacting With Computers* 15(3): 429-452.

Mandryk, R. L., K. M. Inkpen, and T. W. Calvert. 2006. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology* 25(2): 141–158. https://doi.org/10.1080/01449290500331156

Minge, M., & Riedel, L. (2013). meCUE - Ein modularer Fragebogen zur Erfassung des Nutzungserlebens. *Mensch & Computer 2013: Interaktive Vielfalt*.

Minge, M., M. Thüring, I. Wagner, and C. V. Kuhr. 2016. "The meCUE Questionnaire. A Modular Evaluation Tool for Measuring User Experience." In *Advances in Ergonomics Modeling, Usability & Special Populations*, Proceedings of the 7th Applied Human Factors and Ergonomics Society Conference, edited by M. Soares, C. Falcão, and T. Z. Ahram, 115-128. Switzerland: Springer International Press.

Nielsen, J. 1993. Iterative user-interface design. *Computer* 26(11): 32-41.

Petrie, H., and N.BEVAN. 2009. The evaluation of accessibility, usability, and user experience. *The Universal Access Handbook*: 1-16.

Petrie, H., and A.Edwards. 2006. Inclusive design and assistive technology as part of the HCI curriculum. *Proceedings of HCI Educators Workshop'2006*: 23-24.

Petrie, H., F. Hamilton, and N. King. 2004. "Tension, what tension? Website accessibility and visual design." In *Proceedings of the 2004 International Cross-Disciplinary Workshop on Web Accessibility (W4A)*, 13–18. New York, NY: ACM. doi: 10.1145/990657.990660

Petrie, H., and O. Kheir. 2007. "The relationship between accessibility and usability of Websites." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 397–406. New York, NY: ACM. doi: 10.1145/1240624.1240688

Rotondi, A. J., Eack, S. M., Hanusa, B. H., Spring, M. B., and Haas, G. L. 2015. Critical Design Elements of E-Health Applications for Users with Severe Mental Illness: Singular

Focus, Simple Architecture, Prominent Contents, Explicit Navigation, and Inclusive Hyperlinks. *Schizophrenia Bulletin* 41(2), 440–448.

Ruth-Janneck, D. 2011. Experienced barriers in Web applications and their comparison to the WCAG guidelines. In *Information quality in e-health*, edited by A. Holzinger, and K. Simonic, 283-300. Berlin, Germany: Springer. doi: 10.1007/978-3- 642-25364-5_21

Rutter, R., P. H. Lauke, C. Waddell, J. Thatcher, S. L. Henry, B. Lawson, A. Kirkpatrick, C. Heilmann, M. R. Burks, B, Regan, and M. Urban. 2007. *Web accessibility: Web standards and regulatory compliance*. New York, NY: Apress.

Sacharin, V., Schlegel, K., and Scherer, K. R. 2012. Geneva Emotion Wheel rating study (Report). Geneva, Switzerland: University of Geneva. *Swiss Center for Affective Sciences*.

Sarodnick, F., and H. Brau. 2006. *Methoden der Usability Evaluation*. Bern, Switzerland: Verlag Hans Huber.

Sauro, J., & Lewis, J. R. (2009). Correlations Among Prototypical Usability Metrics: Evidence for the Construct of Usability. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1609–1618. https://doi.org/10.1145/1518701.1518947

Sauro, J., & Lewis, J. R. (2012). Standardized usability questionnaires. *Quantifying the user experience*, 185-240.

Sharp, H., Y. Rogers, and J. Preece. 2007. *Interaction design: beyond human-computer interaction*. Wiley: London.

Schmutz, S., A. Sonderegger, and J. Sauer. 2016. Implementing recommendations from web accessibility guidelines: would they also provide benefits to nondisabled users. *Human Factors* 58(4): 611-629. doi: 10.1177/0018720816640962

Schmutz, S., A. Sonderegger, and J. Sauer. 2017. Implementing recommendations from web accessibility guidelines: a comparative study of nondisabled users and users with visual impairments. *Human Factors* 59(6): 956-972. doi: 10.1177/0018720817708397

Schmutz, S., A. Sonderegger, and J. Sauer. 2018. Effects of accessible website design on nondisabled users: age and device as moderating factors. *Ergonomics*, *61*(5), 697-709.

Schmutz, S., A. Sonderegger, and J. Sauer. 2019. Easy-to-read language in disability-friendly websites: effects on nondisabled users. *Applied Ergonomics*, *74*, 97-106.

Shackel, B. 1981. Man-computer Interaction: Human Factors Aspects of Computers & People. In *Proceedings of the NATO Advanced Study Institute on Man-Computer Interaction,* 5-18. Mati, Greece: Springer.

Shneiderman, B. 1998. *Designing the user interface: strategies for effective human-computer interaction.* Addison Wesley Longman: Berkeley.

Sonderegger A., Zbinden, G., Uebelbacher, A. & Sauer J. (2012). The influence of product aesthetics and usability over the course of time: a longitudinal field experiment. *Ergonomics, 55*(7), 713-730.

Sonderegger, A., Uebelbacher, A., and Sauer, J. 2019. The UX Construct – Does the Usage Context Influence the Outcome of User Experience Evaluations? In *Human-Computer Interaction – INTERACT 2019* (pp. 140–157), D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, & P. Zaphiris (Eds.). Springer International Publishing. https://doi.org/10.1007/978-3-030-29390-1_8

Tan, C. S. S., J. Schöning, K. Luyten, and K. Coninx. 2013. Informing intelligent user interfaces by inferring affective states from body postures in ubiquitous computing environments. *Proceedings of the 2013 International Conference on Intelligent User Interfaces - IUI '13*: 235-264.

Thatcher, J., M. R. Burks, C. Heilmann, S. L. Henry, A. Kirkpatrick, P. H. Lauke, . . ., and C. D. Waddell. 2006. *Web accessibility: Web standards and regulatory compliance*. Berkeley, CA: Apress.

Thielsch, M. T., and Thielsch, C. 2018. Depressive symptoms and web user experience. *PeerJ*, 6, e4439.

Thüring, M., and S. Mahlke. 2007. Usability, aesthetics, and emotions in human-technology interaction. *International Journal of Psychology* 42(4): 253-264.

Tractinsky, N. 2018. The Usability Construct: A Dead End? *Human–Computer Interaction* 33: 1-47.

Van Schaik, P., & Ling, J. (2011). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. *Interacting with Computers*, *23*(1), 18-32.

Vermeeren, A. P., J. Kort, A. H. Cremers, and J. Fokker. 2008. Comparing UX Measurements, a case study. *International Workshop*: 72.

Waller, S., M. Bradley, I. Hosking, and P. J. Clarkson. 2015. Making the case for inclusive design. *Applied Ergonomics* 46: 297-303.

Watson, D., L. A. Clark, and A. Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54(6): 1063.

Wright, P. C., J. M. McCarthy, and L. Meekison. 2003. A framework for analysing user experience. *Funology: from usability to user enjoyment.* Dordrecht: Kluwer.

Yesilada, Y., G. Brajnik, and S. Harper. 2009. How much does expertise matter?: a barrier walkthrough study with experts and non-experts. In *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*: 203-210. ACM.

Yesilada, Y., G. Brajnik, M. Vigo, and S. Harper. 2013. Exploring perceptions of Web accessibility: A survey approach. *Behaviour & Information Technology* 34: 119–134. doi: 10.1080/01 44929X.2013.848238