What Interactivity Means to the User
Essential Insights into and a Scale for Perceived Interactivity

Dominik J. Leiner
Institut für Kommunikationswissenschaft
und Medienforschung
Ludwig-Maximilians-Universität München
Oettingenstraße 67
80538 Munich
phone (089) 2180-9401
e-mail: leiner@ifkw.lmu.de

Oliver Quiring
Institut für Kommunikationswissenschaft
und Medienforschung
Ludwig-Maximilians-Universität München
Oettingenstraße 67
80538 Munich
phone (089) 2180-9833
fax (089) 2180-9429
e-mail: quiring@ifkw.lmu.de

Abstract

The selection and use of media depend largely on how users perceive such media. A central aspect of the “new media” is their interactivity, but how users perceive this phenomenon has rarely been researched. This study provides an in-depth investigation into the perception component of interactivity and develops a compact scale for its measurement. According to psychological approaches, practical uses (affordances) – not physical or technical characteristics – guide perception. While existing scales mostly measure whether the “interactive” technical features of devices or websites are noticed, our instrument is based on the affordances that interactivity provides. Consequently, a new research design, the use-identified meaning, was implemented. This is the first study on interactivity that empirically examines a wide range of Internet-based services, thus meeting the broad ambit of interactivity. Our results generally validate the existing constructs, which are largely based on technical characteristics, yet provide additional insights into the relevant contexts and the subjective significance of different aspects of interactivity.
What Interactivity Means to the User
Essential Insights into and a Scale for Perceived Interactivity

Theorists still disagree on the meaning of interactivity (Bucy, 2004; Goertz, 1995; Heeter, 2000; Liu & Shrum, 2002; Wu, 2005). The concept has long been used in “new media” contexts (Schönhagen, 2004), although interactivity is currently understood as a comprehensive construct that is applicable to both traditional mass media and computer-mediated communication (e.g., Bucy, 2004; Heeter, 2000; Jensen, 1998; Quiring & Schweiger, 2006; Steuer, 1992). During the last decade, three perspectives for considering interactivity have emerged in the literature: the (technical) characteristics of media services, the characteristics of a communication process, and the users’ perception (Kiousis, 2002; McMillan, 2000, 2002a). Most contributions to this field are devoted to the technical options (e.g., Durlak, 1987; Goertz, 1995; Jensen, 1998) and few relevant papers take the process perspective (e.g., Rafaeli, 1988; Rafaeli & Sudweeks, 1997). Conversely, it is an established notion that the user’s role has to be considered in order to understand interactivity as a social phenomenon (Krotz, 1995; McMillan, 2000; Vorderer, 2000; Wu, 1999). Given technical options, a user’s behavior is predominantly guided by subjective perception. Consequently, there is a lack of empiric evidence on users’ perception of interactivity as the few existing studies on perception have hardly scrutinized the subjective point of view.

This paper therefore presents empirical research that retraces users’ perspective of interactivity by focusing on “what does interactivity mean to the media user?” Based on the answers to this question, we develop a scale for measuring individually perceived interactivity that supports cross-media application.

2. Research on perceived interactivity

In order to adequately understand a media user’s complex behavior, it is necessary to comprehend the user’s perception (Bouwmann & van de Wijngaert, 2002; Downes & McMillan, 2000; Morrison, 1998; Rodgers & Thorson, 2000; Sohn & Lee, 2005; Wu, 2005). The role of subjective perception becomes particularly evident in the interactive media context: simply using such media requires continuous activity, thus relegating passive consumption to a back

---

1 At least two of these three perspectives are addressed in Wu (2005), Lee, Lee, Kim and Stout (2004), Liu and Shrum (2002), and Stromer-Galley (2004).
seat. Nonetheless, research on the perception of interactivity is easily manageable: one analysis deals with e-mails to an editorial office (Newhagen, Cordes, & Levy, 1995) and three studies are devoted to the meaning that “interactivity” has for different user groups (Downes & McMillan, 2000; Jensen, 2005; Morrison, 1998). Finally, the literature documents three efforts to develop a scale for perceived interactivity (Liu & Shrum, 2002; McMillan & Hwang, 2002; Wu, 1999).

The term “perceived interactivity” first appears in a paper by Newhagen, Cordes, and Levy (1995) who analyze audience responses sent to NBC Nightly News via the Internet in 1993/94. They operationalize perceived interactivity by means of the efficacy that people expected their e-mails to have. They therefore distinguish the “authors’ sense of their own ability to generate an efficacious message,” which is a broadcasting output (“sender interactivity”), from the expectation that a message would be read and answered (“NBC interactivity”). The feeling of being effective and the hope of generating a reaction denote two aspects of interactivity that more current literature calls control and responsiveness (for more extensive current concepts see Jensen, 1998; Kiousis, 2002; McMillan, 2000).

All three studies on users’ notions of “interactivity” by Downes and McMillan (2000), Jensen (2005), and Morrison (1998) rely on qualitative interviews and focus groups; these studies thus directly questioned users on their understanding of the term “interactivity.” The results are barely comparable. Downes and McMillan (2000), who interviewed ten experts, find six dimensions that increase interactivity: two-way communication, flexible timing, a feeling called a “sense of place” (similar to the concepts of “presence” or “social presence,” see Lee, 2004; Short, Williams, & Christie, 1976), the user’s control of the communication process, responsiveness, and the perceived goal of communication. Jensen (2005) conducted eight in-depth interviews with web developers and end users. He identifies three “degrees or varieties” (p. 11) of interactivity: a unique (non-standardized) presentation of contents, other people’s involvement, and self-entertainment on the computer, which includes computer games. These are well-known forms of interactivity rather than interactivity criteria: interaction with a medium or user-to-system interactivity and a medium as a communication tool or user-to-user interactivity (Bucy, 2004; Goertz, 1995; Krotz, 1995; Stromer-Galley, 2004). Morrison (1998) interviewed consumers in respect of entertainment electronics. Her interviewees associated “interactivity” with being active themselves, having control over the communication process, having a feedback channel, and receiving multisensual input (aural, visual, tactile, etc.).
Although each of these studies reveals plausible aspects of interactivity, none clarifies whether the interviewed users really provided information on what they believe interactivity comprises. This doubt is primarily due to the high risk of confusing the marketing term “interactive” with the phenomenon’s subjective meaning. If such confusion were to occur, the interviewees would not mention what interactivity means to them, but merely reproduce core messages from advertising, media coverage, and everyday discussions. The term’s manifold and inconsistent applications may therefore obscure aspects of interactivity that are relevant in respect of users’ behavior.

Studies dealing with the development of a scale only partially map the characteristics stated above, omitting some of them, and amending others. Wu (1999) constructs a scale for perceived interactivity in the light of websites’ marketing effectiveness (Chen & Wells, 1999; Cho & Leckenby, 1999). This scale is solely based on the two factors presented by Newhagen, Cordes, and Levy (1995). User efficacy or control is indicated by a rating of the website’s navigation, while the system’s potential or responsiveness is indicated by the site’s reaction speed. Ten items operationalized these dimensions, seven referring to navigation and three to responsiveness (Appendix 1). The scale validation is predicated by responses to two competing greeting card websites (Wu, 1999). Wu presented a first quantitative measure of perceived interactivity. The applicability of this scale is, however, limited as it was constructed for websites and its dimensions are geared to a concept of interactivity that is now outdated. Additional problems arise from using the website’s navigation as an indicator, as navigation may be confounded with usability (Sundar, 2004, p. 386).

McMillan and Hwang (2002) developed another scale. They collected scale items from the literature, consolidated this item pool through expert interviews (Downes & McMillan, 2000), and thereafter adapted the list of associations with the term “interactivity.” Based on students’ ratings of two commercial websites with different degrees of interactivity, these authors then reduced the resulting item pool of associations to 26 concepts. Finally, they arrived at an 18-item scale (Appendix 2) with three dimensions: real-time conversation, no delay, and engaging. Owing to its focus on reliability, this scale comprises duplicate items that only vary in respect of their valence. Although this overlap formally increases the scale reliability, there is no progress in terms of the content (for methodological criticisms of reworded items see Ajzen, 1988, p. 17). The scale validation is again predicated by responses to just two websites. The main point of criticism, though, is the process of item pool generation: the literature only emphasizes the technical attributes of interactivity (e.g., two-way communication). Such
a scale can indeed measure whether technical attributes are noticed (perceived), but it cannot measure how technology and the communication process, i.e. the experience of interactivity, are perceived. Although superficial attributes are observed, it remains unclear what the user sees behind them (e.g., a synchronous, two-way channel may mean that I am pressed for a quick answer). The latter is, however, the point of interest in perceived interactivity, as it has behavioral implications.

A third scale by Liu and Shrum (2002) is based on an analysis of the scientific concept of interactivity. The authors identified three basic dimensions (active control, reciprocity, and synchronicity) and constructed twelve items for each (Liu & Shrum, 2002). By means of users’ ratings of three existing websites, this initial item pool was consolidated into fifteen items (Liu, 2003; Appendix 3). Further student ratings of two fictitious websites yielded data for a structural analysis. Unsurprisingly, according to the underlying dimensions, three factors fit the items best, although a two-factor model fits the data nearly as well (Liu, 2003). This scale contains multiple items that only formally enhance reliability – and is thus comparable to the scale by McMillan and Hwang (2002). Further, in order to confirm the three factors, it seems infelicitous to determine validation on the basis of three and/or two websites. The major point of criticism is again that the noticing of product attributes is measured (e.g., availability of adequate navigation options and server speed) and not the perception of interactivity.

In summary, each of the scales for perceived interactivity presented hitherto has pragmatic, technical, and/or theoretical shortcomings. First, the scales have been developed for marketing requirements and their items are only applicable to websites. In some cases, the scale construction process is based on focus group discussions of the term interactivity. This could be problematic in the sense that such a construction may primarily return linguistically indicated associations (e.g., being active) and marketing messages. Second, the range of stimulus material used for validation is limited; consequently, the results of the multivariate structural analysis mainly depend on interpersonal variation. Third, in respect of the theoretical level, prior research has been strongly aligned with the aspect of technology. The scales primarily measure the perception of product attributes, not the perception of an interactivity experience. On the whole, the research and scales introduced above provide very limited evidence with which to understand perceived interactivity, its relationship with technical features, and its relationship with the process of lived interactivity. Building on these findings would be ineffective in respect of revealing users’ perspective, as shown in the following section. We therefore searched for another way to approach perceived interactivity.
3. An alternative approach to the operationalization of perceived interactivity

The task of (ap)perception involves a complex processing of stimuli from the environment (bottom-up) and their categorization into the subjectively experienced world (top-down). This process has not yet been considered to ascertain the perception of interactivity. Before a scale for perceived interactivity is constructed, the concept of subjective perception should therefore be traced.

Based on Gestalt psychology, Gibson (1979) develops the ecological approach to visual perception. According to him, humans do not primarily perceive formal attributes of objects (e.g., shape and texture), but “affordances” (options for action) that the objects afford the beholder (Gibson, 1979, p. 127). Bischof (1998) extends this approach beyond visual perception. Based on evolutionary theory, he argues that living creatures inherit a biological “structure” that encodes whether a phenomenon means gain or harm to the individual, therefore (co-)determining how an object is perceived. For example, the chances are good that a snake may be poisonous and do harm, evolution has therefore favored humans who fear snakes and keep away from them. This heritage guides our perception of snakes to this day, which is why when we hear or see a snake, this may lead to an emotional reaction – fear. In our study, however, we are more interested in rational gains and harms (advantages and disadvantages) than in emotions, as measuring the former is usually easier. Summarized, perception follows intersubjective patterns – largely so, at least. Leontiev (2005, p. 50) provides a further differentiation of the concept of meaning. In addition to the ecological-phyllogntical meaning (objective “living meaning”) introduced above, he discusses an ontogenetical “personal meaning.” This personal meaning is created through the “inner activity of meaning-creation” and is based on the use of something for the user (p. 50). Here, Leontiev’s (2005) theory follows learning theory, which states that the functional meaning emerges from the interaction between a person and an object (see Belsky & Most, 1981; Perinat & Sadurni, 1999). Although the term “personal meaning” denotes a wholly subjective perception, we expect consistent patterns here as well. On the one hand, objects have largely consistent uses; consequently, people accumulate similar experiences with similar objects. On the other hand, socialization teaches people various interactions with the environment, e.g., how and for what purpose to use objects.

---

2 Biologically anchored patterns of perceptions and behavior are a well known idea. One of the first to describe this for psychology was C. G. Jung in his comments on the collective subconscious. However, his work focuses on the behavior-guiding aspects of such anchored patterns (Jung, 2000: 46-48).
The approaches introduced hitherto imply a passive individual formed by evolution, experience, and learning. However, the issue of subjective meaning has also been extensively researched in communication science. In this realm, affordances have been labeled “uses and gratifications,” but the central questions remain the same: what does something mean to me? What will I get from it? Within the scope of the uses-and-gratifications approach (Katz, Blumler, & Gurevich, 1974), it has been shown that users regularly comprehend media services as instruments to make life easier (Papacharissi & Rubin, 2000; Rafaeli & Sudweeks, 1997; Schuemie & van der Mast, 1999). The concepts “expected outcomes” (LaRose & Eastin, 2004, p. 361) and “normative images” of communication media (Flanagin & Metzger, 2001; Perse & Courtright, 1993) denote initial efforts to combine the socially transported meaning and the user’s active role. Nevertheless, the gratification scheme lacks an important aspect: disadvantages and non-intended consequences of use (e.g., time consumption, learning efforts, monetary costs, and social disapproval). Uses that constitute an object’s meaning in terms of Gibson (1979) should be understood as a dimension ranging to either side of uses, gratifications, and advantages towards negative side-effects, and disadvantages.

Before measuring meaning as described here, a last issue has to be addressed: uses are not realized until a person utilizes an object, thus interacting with it. Consequently, meaning is not an attribute of an object, but of a situation. Meaning, which guides perception, should therefore be measured in a concrete use context. In keeping with this insight, we elicited answers regarding the utilization of a media service for a given purpose. From the advantages and disadvantages reported by the respondents, we drew conclusions on the (behavioral relevant) meaning of interactivity. Our alternative approach is therefore based on the everyday experience of interactivity, not on the term “interactivity.” This approach avoids biases due to the diffuse meanings of the term, as mentioned earlier.

3.1. Scope of the interactivity concept

Interactivity has always been described as a construct that is applicable to very different situations of media use and communication, including non-mediated communication (Bucy, 2004; Heeter, 1989; Walther, 1996). Nonetheless, prior research limited itself to single media applications, usually websites, and only a few studies researched contexts beyond the Internet (e.g., Morrison, 1998). Consequently, the bulk of the current knowledge concerns website interactivity – a significant shortcoming. In order not to force the concept into a tight corset a priori, the use-contexts we explored comprised different media services. We thus investigated inter-
activity regardless of the specific technical devices, media or applications. This study’s results are therefore applicable to very different situations or objects in the phenomenal environment. Nevertheless, practical considerations disallowed the inclusion of each and every device and media service that has at one time or another been described as “interactive.” The selection of sufficient media services was, however, a critical issue. The meaning realm of what constitutes perceived interactivity would depend on this selection, making it a critical task.

In order to map the scope of interactivity, we first gathered a comprehensive collection of media services associated with the interactivity phenomenon in scientific literature. This list was complemented by searching Google for the term “interaktiv” (German for “interactive”) and scrutinizing the German language online news coverage between January and May 2006 via Google News to include the latest media services and products as well. We thus created an exploratory directory covering about 100 media applications that are commonly regarded as being highly interactive.3 There are at least three independent schemes providing a systemization of the spectrum of interactivity by Cho and Leckenby (1999), McMillan (2000, 2002b), and Oehmichen and Schröter (2004). However, during the classification of existing and supposedly interactive media services into those theoretical categories, consistent parallels emerged between the schemes. For example, each service pertaining to the category “rich content” in one scheme (McMillan, 2000), belonged to the categories “information oriented” (Oehmichen & Schröter, 2004) and “human-message interaction” (Cho & Leckenby, 1999) in the other two schemes. Ultimately, only 6 “domains” were distinguished according to which current media services could be systemized (Figure 1).

By individual information, we understand media services that return customized data after receiving input from the user. An example of this is asking and receiving the weather forecast for one’s vacation resort. In contrast, static information is fully packaged for a mass audience and generally serves as traditional mass media information, like the news at CNN.com. The domain entertainment covers media services whose goals are to entertain, like online games. Practical technical applications do not primarily provide information, but a technical service like downloads and peer-to-peer file sharing services. Transaction services assist the computer-mediated handling of everyday commercial transactions, examples being shopping at “Amazon” or a bank transfer via online banking. Finally, the domain interpersonal com-

3 While we had to use the diffuse term “interactivity” to retrieve potentially interactive services, our investigation of the subjective meaning was fully based on the selected widespread media services (see section 4). We strictly avoided the term itself. To check the degree of interactivity, which is the independent variable in the design, we resorted to theoretical criteria (see section 5.1).
*munication* comprises media services that allow the direct exchange of individual messages between people (e.g., Internet forum or chat, writing an e-mail, or video conferencing).

**Figure 1: Domains of current interactive media services**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Category according to …</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual information</td>
<td>Rich Content</td>
</tr>
<tr>
<td>Static information</td>
<td>Packaged Content</td>
</tr>
<tr>
<td>Entertainment</td>
<td></td>
</tr>
<tr>
<td>Practical technical application</td>
<td>Virtual Transaction</td>
</tr>
<tr>
<td>Transaction</td>
<td>Virtual Community</td>
</tr>
<tr>
<td>Interpersonal Communication</td>
<td></td>
</tr>
</tbody>
</table>

Only three of the six domains have been addressed in the context of perceived interactivity. Newhagen, Cordes, and Levy (1995) analyzed interpersonal communication (e-mails), the interviews conducted by Downes and McMillan (2000) also center on computer-mediated interpersonal communication, while Morrison (1998) scratches the surface of entertainment (home media technology). To date, all efforts to develop a scale for perceived interactivity have, however, been based on the ratings of websites, which mostly hold static information (Liu & Shrum, 2002; McMillan & Hwang, 2002; Wu, 1999).

Although the six domains offer a broad interactivity scope, they are much too abstract to elicit concrete answers from users. Users need to be really familiar with a media service to provide valid evidence regarding its use context. We consequently chose media services from the exploratory directory as domain representatives that are (a) widespread and (b) typical, excluding, for example, online portals that contain services across miscellaneous domains. These services are: 1. weather forecasts on the Internet (individual information), 2. news on the WWW (static information), 3. multiplayer online games (entertainment), 4. music file sharing services (practical technical application), 5. personal e-mail (individual information), 6. online banking, and 7. online shopping (both transaction). While five domains were well represented by widespread typical applications, the domain transaction offered two very dif-
ferent services that are both widespread and have high social relevance. In order not to reduce this domain’s content, both services were included. We therefore investigated seven media services representing six domains.

3.2. Reconstructing subjective meaning

We wanted to investigate interactivity in its use context. In order to measure the meaning of interactivity, we have to understand interactivity as an attribute that is independent of a situation. The degree of interactivity can vary within a broad range: Someone who wants to transmit an individual message may write a letter, but can also use a synchronous instant messenger, or make a phone call. The assumption of an independent attribute is also compatible with the findings in the present literature. Consequently, we can theoretically imagine any use (or communication) situation with another degree of interactivity. This means that if a user compares a highly interactive media service to its less interactive (perfect) counterpart, any difference measured in the meaning must be an outcome of different interactivity.

The respondents can, however, only make a well-grounded comparison of devices if they have experience of using them. Within empiric realization, it is therefore impossible to compare a service to a hypothetic ideal counterpart (only differing in interactivity); consequently, the ideal counterpart had to be replaced with an existing functional alternative for the service under investigation: nobody can compare a normal e-mail to an e-mail that is not interactive, because we cannot visualize such an e-mail. Nevertheless, it is easy to compare an e-mail to a letter. Obviously, there are more differences between such “real” services than just interactivity, which incorporates irrelevant determinants (measurement error) into measurement results. In order to identify the meaning of interactivity from such comparisons, we had to investigate multiple services and ultimately extract those differences to explain interactivity.

Several operationalizations were constructed to measure this difference in meaning. Finally, the best accuracy was obtained in the pretests by simply asking the interviewees what

---

4 Theoretical operationalizations (e.g., Goertz, 1995), as well as experimental studies measure interactivity according to a medium or service’s various attributes (e.g., Bezjian-Avery, Calder & Iacobucci, 1998, p. 24, Sicilia, Ruiz & Munuera, 2005, p. 36, Stromer-Galley, 2004, p. 120-121, Wu, 2005, p. 47). Sundar (2004, p. 385-386) explicitly describes interactivity as an attribute of technology.

5 This is compatible with the current literature as well. Interactivity is no longer regarded as a dichotomous phenomenon (present/absent), but as continuum (ranging from low to high interactivity, e.g., Rogers, 1986; Steuer, 1992; Goertz, 1995).

6 Two pretests were conducted to structure and optimize the questionnaire. The first pretest was dedicated to subscales not discussed in this paper. The second pretest comprised face-to-face pretests with arbitrarily chosen students, using the think-aloud technique. After testing the main study questionnaire for understanding, the participants were also asked to compare the usability of three different measurements in respect of meaning.
they perceived the differences to be. The users were questioned about the advantages and disadvantages of one media service in comparison to another. Such a comparison’s principal benefit is that a concrete base reference is provided. The interviewee does not rate a service according to a rather fuzzy “bad” and “good”, but can rate it as “worse” or “better” than the reference. Individual bias, which occurs if someone is familiar with the latest technology and devices, is thus minimized.

4. Research design and sample

The empiric study includes a (in our case, twofold) preliminary study and a main study (Figure 2). The preliminary study had two goals: first, functional alternatives had to be identified for each media service under investigation (Internet weather forecasts, online news, multiplayer online games, music file sharing, e-mail, online banking, and online shopping) – services that fulfill the same function but differ in their degree of interactivity from the selected media service. Second, use aspects (the advantages and disadvantages) had to be identified in which these services differ from their functional alternatives. By mentioning uses that suggest themselves spontaneously, the interview partners would reveal what, from a subjective perception, constitute the differences between the services (as far as behavior is concerned). This procedure would therefore reveal a use-identified meaning.

To obtain the required data, in-depth interviews were conducted with regular users of the services under investigation. The interviewees were asked what they would do if one of the services were unavailable (e.g., due to technical problems). If they mentioned an alternative, they were then asked to compare the service to this alternative and mention any advantages and disadvantages. In order to ensure comprehensiveness, especially of the use aspects, an additional, open-question survey, using the same question schema, was conducted in the preliminary study. Overall, the guided face-to-face interviews in the preliminary study involved 10 interviewees, who had been deliberately chosen for their very different personal circumstances, to collect as many different uses as possible. The group of interviewees consisted of 4 women and 6 men aged between 23 and 65 years, employed in different sectors and positions, which ranged from a medical student to an agriculturist, graphic designer, middle management employee, and a pensioner. Since using the media services under investigation required Internet access, the study lent itself to utilizing an online survey. The online survey – the sec-

---

differences. The measures that we rejected elicited (a) two absolute ratings for each media service to calculate an arithmetic difference, or (b) one absolute rating and the difference between the services.
ond part of the preliminary study – was announced on 11 Internet bulletin boards and 4 news-
groups, and was completed by 229 participants.

In preparation of the main study, the multitude of use aspects found in the preliminary
study was consolidated into a manageable number of use categories. In order to avoid per-
sonal schemata biases, consolidation was conducted in a focus group of five graduate students
from communication science. Further reduction to even fewer use dimensions would be done
in the main study. In the main study, a second online survey was conducted to quantify the
differences between the media services and their functional alternatives. Our goals were to (a)
elaborate those differences that characterize interactivity across all the services, (b) construct
a scale derived from these differences, and (c) validate this scale.

Figure 2: Research design overview

The preliminary study and the main study were each
split into two parts to complement each other.

The main study survey was announced on 28 Internet bulletin boards and 5 classified
online ad markets from June to July 2006 after obtaining permission from the administrators
and/or operators. In addition, e-mails were sent to a private mailing list and to the e-mail addresses of participants’ friends whose details the participants had provided at the end of the questionnaire. Neither those e-mail-addresses nor personal information related to the data were stored. As a reward for participation, the interviewees could enter a lottery for book coupons. Overall, we received 1021 completed questionnaires that were included in the main study analysis. The participants were aged between 14 to 68 years. The sample was considerably younger than the average German Internet user (69% of the participants were aged between 16 and 28 years), had a surplus of men (75% male), and contained an above-average number of people with a higher education (67% had the equivalent of a high school certificate or university degree). Since we did not intend to measure the representative distribution of the ideas on interactivity, but identify its fundamental meaning, the study’s interest was not jeopardized by the lack of representativeness.

In order to investigate the use dimension relevance and to simultaneously validate the scale, the participants were randomly assigned into two main groups. Each person in the first group (N = 772) was asked to compare one service to one functional alternative, while people in the parallel group (N = 249)7 were asked to rate a media service without reference. The scale’s construction was based on the answers from the first group, while its validation was predicated on the second group’s answers.

5. Results

5.1. Functional alternatives for the services

The guided interviews and the open-ended questions from the first online survey yielded between one and eight functional alternatives for each service under investigation. For each domain, we selected the two alternatives mentioned most often: 1. Internet weather forecasts are often substituted by weather news in television and videotext. 2. If online news were unavailable, most respondents would fall back on newspapers or television. 3. The alternative mentioned for multiplayer online games is most often reading a book, while playing a single-player computer game was mentioned far less. The former was a bit surprising, but was clarified during the guided interviews: online games and reading books similarly serve as pastimes. 4. The most important alternative to writing an e-mail is making a telephone call. Writ-

---

7 Random assignment was double-weighted: First, there were two comparisons for each domain, but only one absolute rating (2:1 weighting). Second, comparisons that were more important for the results received a higher weighting than absolute ratings (additional 3:2 weighting).
ing letters received far less nominations; as the runner-up, it was, however, included in the main study. 5. Instead of banking and shopping online, our respondents would go to a local bank or store. Since we had two services for the transaction domain, one alternative for each was sufficient, others were barely mentioned. 6. In the preliminary study, only a few respondents admitted using music file sharing services. Consequentially, the practical technical domain only yielded commercial online music stores as an alternative. We therefore also included local music stores in the main study. Having twelve alternatives (two per domain) for the main study, we had twelve comparison pairs (a service and its alternative) at hand. The application of this large set of media services and alternatives did not only ensure that all the domains of interactivity were considered during scale development, but was also necessary to avoid random errors through service selection (see section 5.3 for details).

We nonetheless had no evidence that our initial media services were more interactive than their functional alternatives. In order to identify the service from each comparison pair that offers a higher degree of interactivity (the “more interactive” media service), the services in each pair were compared according to seven common criteria of technical interactivity: (a) the users’ typical activity when using a media service and (b) the option range (number of selection and modification options) that is usually used; (c) the theoretically available option range, (d) the options’ effect (e.g., on the content), and (e) the speed, as well as (f) the flexibility in user timing and (g) any apparent system intelligence. The latter can be rated through extraordinary transformation rules (criteria according to Quiring & Schweiger, 2006, p. 13). Such rules allow shopping systems to suggest, for example, other books that might be of interest to their customers. The initial media services generally rate higher along those seven criteria (see Figure 3), which is not surprising, as they were collected as “interactive” services. The phone call is the only exception: from a technical point of view, a phone call provides more interactivity than e-mail due to the faster exchange of messages. As expected from such multi-dimensional comparisons, the more interactive services are not consistently more interactive in each of the criterions. In addition, technical criteria, as applied here, can only provide a heuristic or even superficial measure of the degree of interactivity that a user might experience. Weighting the criteria may improve the predictive value (Koolstra & Bos, 2005), but there is currently a lack of evidence regarding the proper weights to apply. However, this heuristic provides a sufficient a-priori indication with which to judge which media service in each comparison pair probably leads to more interactivity.
After defining 7 representative media services in the six domains (left), the preliminary study identified functional alternatives for these media services (center). Finally, 12 pairs were created to be compared in the main study. In order to (heuristically) identify the more interactive service in each pair, the authors compared the media services to their alternatives according to 7 technical attributes (criteria according to Quiring & Schweiger, 2006, p. 13). A plus (+) indicates that the initial service offers more technical interactivity than its alternative regarding this criterion, a minus (−) indicates the opposite and a dot (·) indicates no differences between the services regarding this criterion.

By and large, the criteria attributes were more distinct for the initial media services. However “writing a personal e-mail” becomes the less interactive service when compared to “making a telephone call.”

### 5.2. Use dimensions

Media service uses were measured according to the advantages and disadvantages of using a media service compared to using a functional alternative. In the preliminary study, a total of 159 different advantages and disadvantages were mentioned in the guided interviews (N = 10) and in the online survey (N = 229). These were different in the sense that the respondents used different words to describe them. There was substantial overlapping regarding the meaning of the words and many advantages described the absence of a disadvantage. It was therefore possible to consolidate the number of use categories to 28 (Appendix 4) without considerable shortcomings regarding the content. In line with the respondents’ original wording, 28 items were derived for a 5-point polarity profile. In the main study, the general usage of the

<table>
<thead>
<tr>
<th>Media service representing a domain of interactivity (typical and widely used service in the domain)</th>
<th>Functional alternative to the media service (identified in the preliminary study)</th>
<th>Technical characteristics of media service compared to alternative</th>
<th>User</th>
<th>System</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWW weather forecast vs. Weather news television</td>
<td>+ +</td>
<td>+ +</td>
<td>+ +</td>
<td>+ ·</td>
</tr>
<tr>
<td>vs. Weather news teletext</td>
<td>− +</td>
<td>+ ·</td>
<td>+ +</td>
<td>− ·</td>
</tr>
<tr>
<td>Online news vs. News in the newspaper</td>
<td>+ +</td>
<td>+ ·</td>
<td>· ·</td>
<td>· ·</td>
</tr>
<tr>
<td>vs. Television news</td>
<td>+ +</td>
<td>+ +</td>
<td>+ ·</td>
<td>− ·</td>
</tr>
<tr>
<td>Multiplayer online games vs. Reading a book</td>
<td>+ +</td>
<td>+ ·</td>
<td>· ·</td>
<td>− +</td>
</tr>
<tr>
<td>vs. Single player games</td>
<td>· +</td>
<td>+ +</td>
<td>· ·</td>
<td>− ·</td>
</tr>
<tr>
<td>Music file sharing service vs. Local music store</td>
<td>+ +</td>
<td>+ +</td>
<td>· +</td>
<td>− +</td>
</tr>
<tr>
<td>vs. Online music stores</td>
<td>· +</td>
<td>+ +</td>
<td>· ·</td>
<td>− ·</td>
</tr>
<tr>
<td>Online banking vs. Go to a local bank</td>
<td>· ·</td>
<td>· ·</td>
<td>+ +</td>
<td>− +</td>
</tr>
<tr>
<td>Online shopping vs. Go to a local store</td>
<td>+ ·</td>
<td>· ·</td>
<td>− +</td>
<td>+ +</td>
</tr>
<tr>
<td>Writing a personal e-mail vs. Making a telephone call</td>
<td>− -</td>
<td>· ·</td>
<td>− +</td>
<td>− -</td>
</tr>
<tr>
<td>vs. Writing a letter</td>
<td>+ ·</td>
<td>· ·</td>
<td>− +</td>
<td>+ +</td>
</tr>
</tbody>
</table>
media services was elicited first. Thereafter one media service that a respondent actually used was chosen at random and a functional alternative added. Each of the 772 interviewees from the first main group therefore compared one service to one alternative on the polarity profile. As some media services are used by more people than others, we used weighted probabilities for random selection. Finally, each of the twelve comparison pairs was rated by 57 to 72 respondents.

As expected, the items within the profile are highly correlated – an effect that can be interpreted as a general preference for one of the two media services compared. To find use dimensions in spite of the strong collinearity, a hierarchical cluster analysis was conducted. We chose 11 clusters after analyzing the cluster distances and consistency regarding the content. This solution’s correlation structure was further scrutinized by using a factor analysis with 11 preset factors and a varimax rotation to find the possible artifacts of the cluster analysis. Three items showed substantial loadings on two factors; as these items were apparently understood ambiguously, they were removed from the analysis. Another item (“more up-to-date”) was so weakly correlated to the other items in the cluster that we had to create an extra use dimension (cluster) from this item. A cluster with only one item was labeled “habitualization”; this cluster was, however, revealed as no more than a description of the use intensity, which was covered by the questionnaire also ($r = 0.7$). This is no genuine advantage and the cluster was therefore not accepted as a use dimension in the subsequent analysis. Overall, we found that users distinguish the media services from the alternatives by 11 use dimensions: price, performance, reliability, speed, effort, time-flexibility, mobility, personal proximity, option range, up-to-dateness, and stimulation (see Appendix 4).

5.3. Characteristics of perceived interactivity

Our comparison pairs are by no means ideal, perfectly matching, functional alternatives. The services and their alternatives do not only differ in their interactivity, but also in respect of other attributes. E-mail, for example, has been described as cheaper than a phone call – an advantage that we do not expect to be linked to interactivity. This problem of irrelevant determinants was solved by identifying those use dimensions that are characteristically linked to more interactive services. The first step so was to check whether the respondents recognized a

---

8 The residual option “not applicable” was replaced by the mean rating for the same service-alternative combination in the cluster analysis. Distance was calculated as the mean Euclidean cluster distance.

9 We excluded “clearer,” “provides more support,” and “yields more detailed information.”

10 The relative use intensity was measured on a 5-point scale ranging from “in comparison, I almost exclusively use A” to “almost exclusively use B”. A and B were replaced by the services compared.
systematic difference between a more interactive service and its less interactive alternative. If the respondents perceived no difference regarding a use dimension, we expected a neutral rating – at the scale center. Consequentially, a significant deviation from the theoretical scale center signals a perceived difference (t-test, p < 0.05, N_{min} = 57, N_{max} = 69, see Figure 4). Time-flexibility, for example, was rated higher for the more interactive services in nine combinations, equal in two, and lower only once (multiplayer online games provide less flexibility than reading a book). Does this mean that time-flexibility is a characteristic of perceived interactivity?

We presumed that the irrelevant determinants – differences between the service and the alternative that cannot be traced back to interactivity – were randomly distributed across the comparison pairs. If a use dimension were therefore irrelevant for interactivity (e.g., price), the interviewees would randomly rate it better in respect of both some of the more interactive media services (e.g., online news vs. newspaper) and less interactive alternatives (e.g., reading a book vs. online games). In the second step, we thus tested the significance of the relationship between the positive and negative ratings. In respect of time-flexibility, there are nine positive and only one negative ratings. Users will therefore generally rate more interactive services as more time-flexible – with an error probability of less than 5% (chi² test, p < 0.05, df = 1). Ultimately, such use dimensions are characteristically perceived in an interactive context.

Up-to-dateness is the only use dimension that received higher ratings in respect of the more interactive service across every comparison. This makes being up-to-date a clear characteristic of perceived interactivity (p < 0.001). Significantly more positive ratings were also found in respect of the use dimensions option range (p < 0.01), time-flexibility, speed, mobility, and minor effort (p < 0.05). Negative ratings regarding those dimensions were notably often found in respect of media services in the entertainment domain (see Figure 4). As expected, the use dimensions price, performance, and reliability are not correlated to the perception of interactivity. Astonishingly, the use dimensions stimulation and personal proximity are also rated ambivalently – they are not necessarily attributes of the more interactive service.

---

11 One could surmise that the item was misunderstood as the technology being more “up-to-date” (org. “aktueller”) than receiving more “up-to-date” information or communication in a situation (as described in the guided interviews). However, the telephone call was rated more up-to-date than writing an e-mail, although the telephone is much older. This is evidence that the respondents understood the item as intended.
Figure 4: Respondents’ comparison of media services vs. their functional alternatives

<table>
<thead>
<tr>
<th></th>
<th>Individual information</th>
<th>Static information</th>
<th>Entertainment</th>
<th>Practical technical application</th>
<th>Transaction</th>
<th>Interpersonal Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>Weather news television</td>
<td>Weather news teletext</td>
<td>Online news</td>
<td>Online news</td>
<td>Online news</td>
<td>Online news</td>
</tr>
<tr>
<td></td>
<td>News in the newspaper</td>
<td>TV news</td>
<td>Internet weather forecast</td>
<td>Interactive services</td>
<td>Interactive services</td>
<td>Interactive services</td>
</tr>
<tr>
<td>Media services compared</td>
<td>Services with higher degree of interactivity (see technical characteristics)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>more interactive functional alternative (reference)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>less interactive functional alternative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aggregate of significant respondents’ ratings p &lt; 0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.s. p &gt; 0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>* p &lt; 0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>** p &lt; 0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*** p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Respondents’ ratings (significant evaluations, p &lt; 0.05)</th>
<th>Up-to-dateness</th>
<th>Options range</th>
<th>Time-flexibility</th>
<th>Speed</th>
<th>Mobility</th>
<th>Effort</th>
<th>Price</th>
<th>Performance</th>
<th>Stimulation</th>
<th>Personal proximity</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>n.s.</td>
<td>+</td>
<td>n.s.</td>
<td>n.s.</td>
<td>+</td>
<td>n.s.</td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Respondents (N = 772) | 60 | 62 | 72 | 63 | 61 | 69 | 64 | 68 | 57 | 69 | 64 | 63 |

1 Making a phone call was found to be the most important alternative to an e-mail. The former is the more interactive service in this pair due to its technical characteristics.

Reading: 60 respondents compared Internet weather forecasts to weather on the television. In this comparison, they rated Internet weather forecasts as cheaper (“price”) but less stimulating (each p < 0.05).

The 11 dimensions of the “respondents’ ratings” (left) were deducted from a cluster analysis of the respondents’ answers to 28 items. By comparing the ratio of significant positive and negative ratings on each dimension, we could distinguish dimensions that are typical of interactive services (above the broken line) from those that are not typical (below the broken line).
In summary, our data reveals the following pattern: up-to-dateness, a broad range of options, fast reactions (speed), time-flexibility, mobility, and effort saving are the practical uses that constitute the meaning of interactivity for the user. Media services in the entertainment domain have a special role that is discussed in the following section. Other advantages have been found in respect of some interactive contexts (price, performance, stimulation, personal proximity, reliability), but they are service-specific and not systematically affiliated to interactivity.

Neither the items nor the use dimensions were provided by the study design. Nor did the research instruments address the term “interactivity.” Nevertheless, the use dimensions found in our study broadly match the existing concepts of the construct interactivity. Several technical characteristics (e.g., speed, options) that are indicators of interactivity are also perceived as uses of interactive services. By and large, our findings can be regarded as a validation of the existing concepts’ content. However, there are technical features of interactivity (e.g., activity, two-way-communication) that do not seem to feature in the users’ perception. Conversely, mobile computing is a new development, but already seems to influence the users’ idea of interactivity. This study also adds the use aspect up-to-dateness, which has not as yet been considered in literature, although the items used for its measurement may partly confound up-to-date information/communication with more current services. Further, our respondents did not perceive stimulation and personal proximity as necessarily associated with interactivity, although varieties of those characteristics are thought to be important facets of interactivity (e.g., Bretz, 1983; Coyle & Thorson, 2001; Durlak, 1987; Massey & Levy, 1999).

5.4. Computers and people are perceived as decisively different

The respondents’ ratings have very clearly shown that multiplayer online games (entertainment domain) and making a phone call (interpersonal communication) are perceived very differently than services in other domains are. Although these services provide many “interactive” technical characteristics, the respondents in our study rarely perceived them as more interactive and partly even perceived them as less so (see Figure 4). Conversely, these services are rated high in respect of stimulation and personal proximity. The latter is not surprising, because each of the three services requires substantial interaction with another person (the conversational partner, a team mate or opponent). Vice versa, interviewees often indicated that personal proximity was “not applicable” (the residual option, also see section 6) with re-
gard to other services like information and transaction services where no person is “visible at the other side.”

These observations, first, point out a limitation of the research design employed here. When searching for the characteristic attributes of a range of objects (different services in our case), one is inevitably limited to central and across-the-board components. Further aspects, valid for a part of the services only, will be deliberately ignored. We expected that each service would, by and large, be perceived as more interactive regarding the central components. Multiplayer online games, however, which are often described as a prime example of interactivity, were systematically found not to conform. One interpretation of this could be to doubt the implicit assumption that people have a general perception of a prototype “interactivity” – an interpretation which is strongly at variance with the assumptions in the existing literature. Another interpretation seems more likely: we assume that not all services that are interactive in the technical sense are perceived as interactive in the sense of the central component. There may be two (or more) dimensions of perceived interactivity, or general perception prototypes that are both called “interactivity.” In our study, one of them comprises the advantages that we identified as typical of interactivity. The other one may comprise personal proximity and stimulation, although we do not have appropriate data to further scrutinize this assumption.

This finally leads to the well-known and often broached separation between user-to-user and user-to-system interactivity (Bucy, 2004; Goertz, 1995; Krotz, 1995; Stromer-Galley, 2004) and confirms this duality in respect of the perception level. Media users seem to clearly distinguish both these constructs. While user-to-user interactivity is essentially interlinked with social exchange processes, user-to-system interactivity (Cho & Leckenby, 1999; Quiring & Schweiger, 2006) involves a distinctively individual direction. The latter might be tightly interlinked with concepts like usability, comfort, or practical use value. The meaning of interactivity that we found through this study can also be classified into user-to-system interactivity. Although this meaning cannot represent the full range of the interactivity construct, the aspect of perceived user-to-system interactivity is highly relevant. Not only does the major share of interactive services belong to the user-to-system type, but, as this study confirms, this aspect of interactivity also involves the bulk of (potential behaviorally relevant) use dimensions.

Once again, our results reflect the necessity to neatly distinguish between interactivity and (social) interaction, as they share similar terms. Once again, we have doubts about the
theoretical unit “interactivity,” as it may comprise two constructs that are too different to be pragmatically used simultaneously.

6. A scale for the parsimonious measurement of perceived user-to-system interactivity

The items that identify the subjective meaning of interactivity are self-evidently well suited to also measure perceived interactivity. Comparing two media services in a polarity profile does not, however, seem feasible for measuring interactivity perceived in respect of one service. In our parallel group in the main study (N = 249), we therefore asked the respondents for non-comparing ratings of a service’s attributes on a 6-point Likert scale\(^\text{12}\). The ratings for each service along the 11 use dimensions were nevertheless very similar to the comparing ratings. Figure 5 exemplifies this in respect of news on the Internet. We assume that when asked for an absolute rating, the respondents implicitly make a comparison with some functional alternative. This has to be considered in the interpretation of measurements since functional alternatives are likely to change and shift over time.

Figure 5: Similar non-comparing and comparing ratings for online news

```
Each service was rated along all 28 items. As we are only interested in the 6 use dimensions that constitute perceived user-to-system interactivity, only 15 of those items\(^\text{13}\) remained relevant for the scale construction. Given our goal of a parsimonious measure, we identified one

---
\(^{12}\) We used a 6-point scale in keeping with the German grading system. The scale ranges from “does not apply at all” to “applies fully”.

\(^{13}\) “is time-flexible, is usually at hand” (time-flexibility), “makes few demands on me, requires little effort, takes me little time, is comfortable” (effort), “causes no delay, means little waiting time, is fast” (speed), “can be used anywhere” (mobility), “is up-to-date” (Up-to-dateness), “is able, is complete, is versatile, is individual” (broad option range)
item for each use dimension\(^{14}\) that (a) supports the scale’s reliability and (b) allows a valid measurement. Reliability was indicated by the inter-item correlation, which was calculated by means of the item and the full scale of 15 items. Validity drew on the residual option “not applicable,” which could be selected instead of a rating (Figure 6). By choosing this option, a respondent indicated that he or she did not see a strong coherence between an item and the object. As such items would lead to invalid ratings (nonattitudes), the reciprocal selection percentage of this residual option was regarded as a validity indicator.

Figure 6: Residual option within the questionnaire

<table>
<thead>
<tr>
<th>Writing a personal e-mail</th>
<th>does not apply at all</th>
<th>applies fully</th>
<th>not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>... requires little effort</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
</tr>
<tr>
<td>... provides direct communication</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
</tr>
<tr>
<td>... is sociable</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
</tr>
<tr>
<td>... causes no delay</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
</tr>
<tr>
<td>... is versatile</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
<td>○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○</td>
</tr>
</tbody>
</table>

An explicit “don’t know” option was provided for each rating item to determine nonattitudes. In this type of question, nonattitudes are most likely to occur when an interviewee has never thought about the object in the sense of the given item. The residual option was therefore labeled “not applicable” (original German: “passt nicht”). This figure is a cutout from the non-comparing rating questions, translated from German.

The resulting shortened 6-item scale (Figure 7) achieves a reliability coefficient of a Cronbach \(\alpha = 0.70\) (\(N = 207\)), which is acceptable given the broad semantic field that is covered by only a few items. A check was done that the short scale adequately maps the original full scale and yields acceptable results: The short scale explains the full scale’s variation (15 items user-to-system interactivity) by 89% (\(r = .95\), \(N = 222\)). As the short scale requires only 6 of 15 items, this truncation seems very reasonable in a parsimonious sense.

\(^{14}\) Simple item consolidation or, better, reduction by inter-item correlation is susceptible to shortcomings regarding the content. Reliability in the sense of uni-dimensionality was therefore balanced against completeness.
Figure 7: Short scale for measuring perceived user-to-system interactivity

<table>
<thead>
<tr>
<th>Use dimension</th>
<th>Items: “The service...”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-to-dateness</td>
<td>... is up-to-date</td>
</tr>
<tr>
<td>Time-flexibility</td>
<td>... is usually at hand</td>
</tr>
<tr>
<td>Speed</td>
<td>... is fast</td>
</tr>
<tr>
<td>Mobility</td>
<td>... can be used anywhere</td>
</tr>
<tr>
<td>Option range</td>
<td>... is versatile</td>
</tr>
<tr>
<td>Effort</td>
<td>... requires little effort</td>
</tr>
</tbody>
</table>

After identifying six use dimensions (left) that are systematically perceived in more interactive contexts, we identified one item (right) for each dimension that provided optimal validity and reliability (based on both criteria’s mean rank), thus forming a 6-item scale for perceived user-to-system interactivity.

7. Discussion

In the past decade, three perspectives have emerged on the construct of interactivity: technical attributes that allow interactivity, the users’ perception, and the resulting use process. However, their relationship has not been well scrutinized to date, not least due to the lack of a scale that can independently operationalize perceived interactivity without reference to its technical attributes. The scales presented hitherto have been geared to the term and the theoretical concept of interactivity, neglecting the subjective perspective, which is linked to a complex apperception process. Furthermore, these scales’ validations have been based on very few websites.

The goal of our study was therefore to develop a parsimonious scale for measuring perceived interactivity that (a) does not depend on technical attributes and the marketing-term “interactivity” and (b) is equally applicable to different domains of interactivity. Consequently, we employed the “use-identified meaning” as an alternative research design. A reconstruction of the subjective meaning of the phenomenon interactivity preceded the scale construction process. This design is intended to incorporate much more information from the respondents – the users – into the resulting research instrument than classic scale development does. Furthermore, this study is a first effort to include a broad range of interactive services. This not only circumvents the diffuse meanings of the term interactivity, but also fulfills the phenomenon’s requirement in respect of its cross-media applicability.

At first glance, the resulting scale seems to lack substantial aspects of the interactivity construct as established in the literature. Neither feedback nor responsiveness or control over
the communication process is directly addressed in the scale’s items. Given the approach that we employed, this was, however, never expected. The item construction is mostly based on statements from “the man in the street,” who usually does not employ the exact phrases that academics use for partial aspects of the concept. This is, of course, no fundamental disadvantage: on the contrary, this step away from scientific wording to behaviorally relevant criteria that users experience directly is the strength of this scale. The scale’s anchorage in the subjective world allows perception to be measured independently without reference to technical attributes. The scale therefore offers empiric access to the relationship between interactive characteristics and perceived interactivity. In addition, the cross-media applicability allows comparisons of interactivity between different communication modes and/or media (e.g., e-mail versus SMS) or devices in respect of perception.

A second glance reveals surprising congruence between this empirically built scale and theoretical concepts of interactivity. The item “fast,” for example, is likely to represent the user-relevant manifestation of responsiveness, while the items “is versatile” and “requires little effort” essentially describe control over the communication process. With the exception of up-to-dateness, each characteristic of perceived interactivity included on the scale can be found in common theoretical concepts (Kiousis, 2002; McMillan & Hwang, 2002). Consequently, the study at hand at least partly validates the existing scales of perceived interactivity that did not necessarily fully capture the construct.\(^\text{15}\)

In summary, the high degree of congruence between the theoretical concepts and our empirical findings, which are based on user statements and ratings, argues for a relatively stable concept of interactivity. Our results moreover provide evidence of the relative importance of each aspect regarding behavioral relevance. Two well-known, basic forms of interactivity – user-to-system and user-to-user interactivity – have been observed in the users’ perception as well. Users clearly seem to distinguish these forms more fundamentally than most theoretical concepts do, that propagate an integrative interactivity.

Nonetheless, our efforts clearly failed to construct a scale for perceived user-to-user interactivity, which is attributable to our domain selection and the focus on practical (not, e.g., social) uses. Our data of non-responses (the residual option “not applicable”) suggests, however, that such a scale would validly only apply to services that visibly involve multiple peo-

\(^{15}\) Some personality attributes were collected in the main study in order to test the criterion validity. We could confirm a medium correlation between need for cognition and perceived interactivity (measured with a 4-item short scale) as found by Sohn and Lee (2005), but with the coefficients varying significantly between the services. We could also confirm a significant correlation between a person’s computer skills and perceived interactivity (Jee & Lee, 2002), but only for the domains transaction and individual information.
ple in the communication process – like in multiplayer online games or e-mail. Conversely, the presented scale for user-to-system interactivity is generally applicable to any media and even other objects that provide (or do not provide) interactivity in any of the various domains. This scale’s application admittedly demands attention to technical development effects. As users seem to keep functional alternatives in mind even when “absolutely” rating a service, the perceived user-to-system interactivity measured with this scale will change over time: as soon as new services fulfill the same tasks even better, old services will be rated by new measures.

References


Appendix 1: Items for measuring perceived interactivity by Wu (1999, p. 11)

1. While I was on the site, I was always aware where I was
2. While I was on the site, I always knew where I was going
3. While I was on the site, I was always able to go where I thought I was going
4. The hyper-linked images and texts tell me exactly what to expect
5. The visual layout was like a roadmap during my exploration of the site
6. When I clicked on hyper-linked images or texts, I felt good about the instantaneous display of information
7. While I was on the site, I could quickly jump from one page to another
8. I felt I did not get much useful information simply because it had too much information
9. I was delighted to be able to choose which link and when to click
10. I was pleased to express my feelings and opinions on the spot through email or feedback form


Real-Time-Conversation
1. Enables two-way communication
2. Enables concurrent communication
3. Non concurrent communication
4. Is interactive
5. Primarily one-way communication
6. Is interpersonal
7. Enables conversation

Engaging
11. Variety of content
12. Keeps my attention
13. Easy to find my way through the site
14. Unmanageable
15. Doesn't keep my attention
16. Passive
17. Immediate answers to questions
18. Lacks content

No Delay
8. Loads fast
9. Loads slow
10. Operates at high speed
Appendix 3: Scale to measure the interactivity of websites by Liu (2003, p. 210)

Active control
1. I felt that I had a lot of control over my visiting experiences at this website
2. While I was on the website, I could choose freely what I wanted to see
3. While surfing the website, I had absolutely no control over what I can do on the site
4. While surfing the website, my actions decided the kind of experiences I got

Two-way communication
5. The website is effective in gathering visitors’ feedback
6. This website facilitates two-way communication between the visitors and the site
7. It is difficult to offer feedback to the website
8. The website makes me feel it wants to listen to its visitors
9. The website does not at all encourage visitors to talk back.*
10. The website gives visitors the opportunity to talk back

Synchronicity
11. The website processed my input very quickly
12. Getting information from the website is very fast
13. I was able to obtain the information I want without any delay
14. When I clicked on the links, I felt I was getting instantaneous information
15. The website was very slow in responding to my requests
Appendix 4: Overview of the use dimensions and associated items

<table>
<thead>
<tr>
<th>Use dimension (number of items)</th>
<th>5-point polarity profile for comparing rating</th>
<th>Non-comparing 6-point scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;Compared to service A Service B …&quot;</td>
<td>&quot;The service …&quot;</td>
</tr>
<tr>
<td>Price (1)</td>
<td>is more expensive is cheaper is cheap</td>
<td></td>
</tr>
<tr>
<td>Performance (2)</td>
<td>provides a poor performance/quality provides a better performance/quality provides a good performance/quality</td>
<td></td>
</tr>
<tr>
<td></td>
<td>provides a poorer presentation provides a better presentation provides a good presentation</td>
<td></td>
</tr>
<tr>
<td>Reliability (2)</td>
<td>is less dependable is more dependable is dependable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is less risky is less risky involves few risks</td>
<td></td>
</tr>
<tr>
<td>Personal proximity (3)</td>
<td>is less personal is more personal is personal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is less sociable is more sociable is sociable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>provides less direct communication provides more direct communication provides direct communication</td>
<td></td>
</tr>
<tr>
<td>Speed (3)</td>
<td>causes delays causes less delay causes no delay</td>
<td></td>
</tr>
<tr>
<td></td>
<td>means more waiting time means less waiting time means little waiting time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is slower is faster is fast</td>
<td></td>
</tr>
<tr>
<td>Effort (4)</td>
<td>makes many demands on me makes fewer demands on me makes few demands on me</td>
<td></td>
</tr>
<tr>
<td></td>
<td>requires more effort requires less effort requires little effort</td>
<td></td>
</tr>
<tr>
<td></td>
<td>takes me more time takes me less time takes me little time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is less comfortable is more comfortable is comfortable</td>
<td></td>
</tr>
<tr>
<td>Time-flexibility (2)</td>
<td>is less time-flexible is more time-flexible is time-flexible</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is at hand less often is at hand more often is usually at hand</td>
<td></td>
</tr>
<tr>
<td>Up-to-dateness (1)</td>
<td>is less up-to-date is more up-to-date is up-to-date</td>
<td></td>
</tr>
<tr>
<td>Mobility (1)</td>
<td>the location is more restricted can be used at more places can be used anywhere</td>
<td></td>
</tr>
<tr>
<td>Manifold option range (4)</td>
<td>can do less can do more can do everything</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is less complete is more complete is complete</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is less versatile is more versatile is versatile</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is less individual is more individual is individual</td>
<td></td>
</tr>
<tr>
<td>Stimulation (1)</td>
<td>is less exciting is more exciting is exciting</td>
<td></td>
</tr>
<tr>
<td>Items excluded from analysis (4)</td>
<td>I am not very used to I am more used to I am used to</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is more confusing is clearer is clear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>provides less support provides more support provides good support</td>
<td></td>
</tr>
<tr>
<td></td>
<td>yields less detailed information yields more detailed information yields detailed information</td>
<td></td>
</tr>
</tbody>
</table>

The development of the 28 items was based on responses to the following question in the preliminary study: “Which advantages does service A yield compared to [the alternative named by the respondent]?” The items were categorized into 11 use dimensions (left) using a hierarchical cluster analysis.