The Experience of Presence: Factor Analytic Insights

Abstract

Within an embodied cognition framework, it is argued that presence in a virtual environment (VE) develops from the construction of a spatial-functional mental model of the VE. Two cognitive processes lead to this model: the representation of bodily actions as possible actions in the VE, and the suppression of incompatible sensory input. It is hypothesized that the conscious sense of presence reflects these two components as spatial presence and involvement. This prediction was confirmed in two studies ($N = 246$ and $N = 296$) assessing self-reports of presence and immersion experiences. Additionally, judgments of “realness” were observed as a third presence component. A second-order factor analysis showed a distinction between presence, immersion, and interaction factors. Building on these results, a thirteen-item presence scale consisting of three independent components was developed and verified using confirmatory factor analyses across the two studies.

Presence is a construct, a variable with various levels and dimensions.

Biocca and Delaney (1995, p. 62)

1 Introduction

When we work or play within virtual environments (VEs), travel through them and interact with virtual objects, it is common that a certain sense of being in the virtual environment, or presence, develops. Except for cinema, where it is known as the diegetic effect (Burch, 1979), this experience is not that common in traditional media. In contrast, interactive media that present a three-dimensional space for the user, such as virtual reality and 3-D games, seem to be a reliable source of this experience. An example can illustrate this. When we read an article about a narrow suspension bridge, we would rarely experience any sensations because of the mentioned height, but we have a clear mental model of the described space. When we see the bridge in an action movie and we look down to the bottom of the valley together with the endangered protagonist, it is likely that we feel fear because of the height. However, when users have to walk over that bridge in a virtual environment, many of them will experience physiological symptoms and sensations of fear, because they have a sense of actually being there (Regenbrecht, Schubert, & Friedmann, 1998).

In this paper, we argue that all three examples basically build on the same cognitive processes. We will start with an analysis of the cognitive processes that lead to the emergence of presence. We will then show empirically that these cognitive processes surface in subjective experiences of presence.
2 Cognition as a Mediator

It is now widely acknowledged that presence should be treated as a psychological phenomenon: “Presence is a state of consciousness, the (psychological) sense of being in the virtual environment.” (Slater and Wilbur, 1997, p. 605). In defining presence psychologically, Slater and colleagues (for example, Slater, 1999; Slater & Usoh, 1994) distinguished it from the concept of immersion, which they defined as an objective description of the technology, describing “the extent to which the computer displays are capable of delivering an illusion of reality to the senses of a human participant.” (Slater & Wilbur, 1997, p. 604f). Although immersion is objectively quantifiable, presence—or, more precisely, the sense of presence—is a subjective experience and only quantifiable by the user experiencing it.

In current theoretical models, the sense of presence is seen as the outcome or a direct function of immersion. The more inclusive, extensive, surrounding, and vivid the VE is (Slater & Wilbur, 1997), or the more similar the transformations in the VE are to those in the real world (Barfield & Hendrix, 1995; Bystrom, Barfield, & Hendrix, 1999), the higher the presence. Person variables are only considered insofar as they moderate the impact of different immersion variables, and, when they are considered, the focus is on stable personality properties (Slater, 1999).

It would be misleading, however, to assume a one-to-one relationship between immersion and presence. One must take into account the cognitive processes leading from stimuli perception to presence. Cognitive processes mediate the impact of immersion on the development of presence. This is different from saying that person variables moderate the influence of immersion (Baron & Kenny, 1986). Stimuli from a VE are only the raw material for the mind that constructs a mental picture of a surrounding world, instead of a mental picture of pixels on the display in front of the eyes (which would be equally valid, or even more so). The sense of “physical reality” is “a consequence of internal processing rather than being something that is developed only from the immediate sensory information we receive” (Ellis, 1991, p. 874). Only by taking these processes into account is it possible to understand how 3-D games, played on a monitor and without direct mapping between body movements and corresponding movements in the virtual world, and even text-only VEs can elicit high degrees of presence (Schiano, 1999; Towell & Towell, 1997).

3 Presence and Mental Models: Construction and Suppression

When users are present in a VE, the outcome of the cognitive processes can be conceptualized as a special type of mental model of the virtual space, in which the location of the own body is construed as being contained in the space rather than looking at it from outside (Biocca, 1997, Regenbrecht et al., 1998). Theories of cognition and memory in the embodied cognition framework (Lakoff, 1987; Lakoff & Johnson, 1999) offer interesting possibilities for understanding the nature of this mental model and the cognitive processes that lead to it. These theories have the further advantage that they explain both the understanding of symbols and of natural environments in one framework.

Consider the following possibility: you mentally represented a VE in terms of what you, with your body, can do in it. Is it possible that presence then depends on which actions you consider possible in the VE? Adopting the framework laid out by Glenberg (1997), we argue that a virtual environment, like every other environment, is perceived and understood by mentally combining potential patterns of actions. Following Glenberg, we call this process the construction of meshed sets of patterns of actions. These patterns represent possible actions in the virtual space: understanding the world means conceptualizing it in terms of actions, or, as Zahorik and Jenison (1998) put it, “presence is tied to action in the environment” (p. 80).

Two sources of patterns of actions exist: projections from the environments and memory (Fodor & Pylyshyn, 1981). First, actions may be directly afforded by the VE when objects follow bodily constraints (Gibson, 1979); these properties are called projectable properties. A bridge over a precipice may have the direct affordance
of walking over it, whether in a real or a virtual environment. But, secondly, patterns of actions from memory can also be meshed with other patterns. This is the function of implicit memory (Glenberg, 1997) and reflects the integration of nonprojectable properties. Thus, we may remember that we have been warned of the bridge because it may break. Importantly, understanding mental models of VEs as embodied models implies that they are not analogous depictions similar to the 3-D model in the computer, but represent actions in this 3-D space. Thus, they are spatial-functional models, not purely spatial models (cf. Glenberg, 1997; but see Zahorik & Jenison, 1998).

Of course, users who are not present in a VE also may have a mental model of the stimuli and the interface. In this case, they just interact with a display or with a space behind the display. For instance, a programmer may focus on the rendering and shading when looking at the displayed VE, without paying attention to his or her position in the space. However, when presence emerges, the structure of the mental model changes dramatically. The actions that are then represented mentally are bodily actions within the space depicted, being functionally related to navigation, manipulation of objects, or interaction with other agents. It becomes important how the space is related to the body. Embodied presence develops from the mental representation of navigation (movement) of the body (or body parts) as a possible action in the virtual world (Schubert, Friedmann, & Regenbrecht, 1999).

Constructing mental models of VEs is similar to both conscious recollection and language comprehension in the sense that it is necessary to mentally represent a situation that is different from the one immediately present. Glenberg (1997) argues that to be able to do this, we have developed the ability to suppress immediate sensory input. To understand the written message about a broken bridge, we have to suppress the immediate sensory input from our surrounding environment, which is normally “clamped” to assure attention to the real environment. Similarly, when perceiving a VE, usually a number of conflicting sensory inputs must be suppressed, including distracting stimuli from the hardware or the real environment. In short, the suppression of conflicting stimuli and the allocation of attention to virtual stimuli (Bystrom et al., 1999) are necessary conditions for presence.

Construction of the spatial-functional model from VE stimuli and suppression of incompatible stimuli from the real world go hand in hand. Both processes are active processes, which must be learned and may break down under overload. In our laboratory, we frequently observe that users new to the hardware and software first have to learn how their actions change the pictures seen in the head-mounted display (HMD), and how to make sense of them (Regenbrecht, 1999). As experts in VE perception, we may have forgotten that it is sometimes hard work to be present.

Both construction and suppression show how closely immersion and the mediating cognitive processes are related. Construction can be supported by the VE, for example, by appropriate depth cues. Suppression can be facilitated or hindered by the hardware and environmental conditions. Taking into account both immersion and cognition enables us to explain differences between not only technologies, but also between people and between situations.

4 Components of the Presence Experience

It is doubtful whether the sense of presence is a unitary experience (Biocca & Delaney, 1995; Kalawsky, 1998; Sheridan, 1992, 1996; Welch, Blackmon, Liu, Mellers, & Stark, 1996). The cognitive analysis above may help to characterize its elements, but first the relation between cognitive processes and subjective experience has to be explained. Conscious experiences are never direct representations of our cognitive processes, as one might assume. In contrast, just as experiences of the environment are highly selective and modified interpretations, they are highly selective and metaphorically modified interpretations of our own cognitive representations and processes (Prinz, 1999). The sense of presence is a conscious experience. We propose that a presence experience (the sense of presence) results from the interpretation of the mental model of the VE, which is
the outcome of the cognitive processes. We become consciously present as an effect of interpreting our own mental construct.\footnote{This distinction is also important for theorizing about consequences of presence. Two examples of reported correlates of the sense of presence are direct motor effects, taken to be objective measures of presence (Freeman, Avons, Meddis, Pearson, \& IJsselsteijn, 2000), and emotional responses such as phobic fear of virtual stimuli (Carlin et al., 1997; Regenbrecht et al., 1998; Rothbaum et al., 1995). The presence measures taken in typical studies on these topics are self-reports and thus tap our interpretations of our own cognition. The researched variables, however, might depend on the cognitive processes rather than on the experience. Because the latter depends on the former, experience and outcomes may also correlate, but not always perfectly so (Welch et al., 1996).}

We argued that two cognitive processes are involved in the emergence of presence: construction of a mental model and attention allocation. Conscious presence experiences should reflect these two processes: presence should involve awareness of possible action patterns and the awareness of the attention allocation necessary to construct it. Therefore, the sense of presence should involve at least two components: the sense that we are located in and act from within the VE, and the sense that we are concentrating on the VE and ignoring the real environment.

This proposal is consistent with Witmer and Singer (1998), who also introduce a distinction between an attention side (involvement) and a spatial cognitive side (psychological immersion) of presence (albeit not on theoretical grounds).\footnote{Witmer and Singer call the spatial cognitive component immersion, but see it clearly as a subjective experience (“the perception of being enveloped”, p. 225). To avoid confusion, we will call their concept here psychological immersion instead of using their term immersion.}

Involvement is a psychological state experienced as a consequence of focusing one’s energy and attention on a coherent set of stimuli or meaningful related activities and events. . . . [Psychological] immersion is a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences. (p. 227)

Witmer and Singer (1998) present cluster-analyzed data from a survey study, but they were not able to show this dichotomy in their data. This is probably due to the low number of items directly assessing subjective presence experiences (cf. Slater, 1999). Instead, Witmer and Singer (1998) aimed at including items that “address factors that influence involvement as well as those that affect [psychological] immersion.” (p. 228, italics added; cf. Singer \& Witmer, 1999). In contrast to Witmer and Singer’s approach, the main goal of the present research is to explore components within the presence construct, and the focus is on the distinction between spatial-constructive and attention components. We hypothesize that this distinction should manifest itself in self-reports: items measuring spatial presence should correlate highly with each other and less with items measuring involvement. Thus, they should be identifiable as different factors in a factor analysis.

Furthermore, items designed to assess the sense of presence should load on other factors than those items that tap subjective evaluations of the immersive technology. To explore whether this distinction shows up in empirical data is a second goal of this research. As a third goal (but not the central focus of this paper), we explore possible distinctions inside the subjective immersion evaluations (Steuer, 1992; Witmer \& Singer, 1998; but see Barfield \& Weghorst, 1993).

We will present two studies that both use exploratory factor analyses to approach these questions. We will then refine these analyses in a third step, computing confirmatory factor analyses to reach a parsimonious presence scale.

5 Study 1

To address the goals outlined above, we conducted a survey study on presence and immersion experiences. The survey combined questions from previously published questionnaires, and from our own past research with newly designed questions. Study 1 was conducted to collect data on both immersion and presence experiences and to explore possible factors inside the concepts.
5.1 Materials and Procedure

In the questionnaire, we included the complete presence questionnaire by Witmer and Singer (1994), as well as items from Ellis, Dorighi, Menges, Adelstein, and Jacoby (1997), Carlin, Hoffman, and Weghorst (1997), Hendrix (1994), Slater, Usoh, and Steed (1994), Towell and Towell (1997), and Regenbrecht et al. (1998). English questions were translated into German\textsuperscript{3} and combined in one 75-item survey. Original scale anchors were kept, but all items were transformed into five-point scales. New items were mostly Likert-type items, anchored with \textit{not at all true} and \textit{completely true}. Additional items assessed the technical equipment with respect to type and quality of the visual and audio output devices, interaction devices, type of software application, duration of use, and number of and potential interactions with other users in the VE. Gender and age were also assessed.

Users of all forms of VE (users of VR or CAVE-like systems, desktop VR, and text-based VE systems, and players of 3-D games) were asked to complete the questionnaire, which was posted on the Web (Müller, 1998). The participants were instructed to remember one of the last times they used a VE and to answer all questions with reference to that single episode only. Advertisements were posted on the Web in various newsgroups and on the site of a German computer game journal. To receive debriefing emails explaining the purpose and results of the study, participants had to provide an email address. Care was taken to provide anonymity.

5.2 Participants

Data are available for 246 participants, some of whom were surveyed after using the VR equipment (HMD-based) in our own laboratory. The other participants answered the questionnaire on the Web. Approximately 10\% of the participants were female, and 90\% male. (No exact numbers can be given due to missing values.) The mean age was 24.5 years, with a standard deviation of 5.3, ranging from 10 to 50 years. The majority (\( n = 224 \)) used monitors, and only 19 experienced the VE via HMD or a system with multiple projections, such as CAVEs (missing data for three cases). Of the monitor users, 195 heard stereo sound and 10 heard mono sound. The main type of application were 3-D games with a first-person perspective (\( n = 191 \)), in contrast to lower percentages of visualization and walk-through applications (\( n = 24 \)) and other games (\( n = 31 \)). Interestingly, 142 users (57.7\%) shared the VE with other users.

5.3 Results

5.3.1 Sample Adequacy. Factor analyses demand high numbers of participants and high interitem correlations for reliable results. Different criteria assured the usability of our sample and the interpretability of the results: with 246 participants, the minimum \( N \) recommended by Guilford (1956) was met, who argued that the minimal sample size should be 200. Two ratios—of subjects to variables and of subjects to factors—are important. The first was 246:68 (3.62), well above the ratio of 2:1 advised by Kline (1994). For the following analysis, the second ratio equaled 246:8 (30.75), the minimum recommended ratio being 20:1 (Arrindell & Ende, 1985). Guadagnoli and Velicer’s (1988) factor stability (FS) index, computed out of sample size and lowest interpreted loading (which is 0.458 in our case), equaled 0.919, indicating a good fit between “true” and sample factor structure.\textsuperscript{4} As a last criterion, the Kaiser-Meyer-Olkin (KMO) measure of 0.827 indicated a very high sampling adequacy and good preconditions for factor analyses (Kaiser & Rice, 1974).

5.3.2 Factorization. The data were factorized using principal components analysis and rotated using oblique direct Oblimin rotation (\( \delta = 0 \)). Oblique rotation was preferred over orthogonal rotation because the factors could not theoretically be assumed to be in-

\textsuperscript{3} In the following text, original English items with references to the authors are used where available. New German items were translated into English for this paper, but they have not been validated empirically so far.

\textsuperscript{4} Factor structures with FS < 0.8 should not be interpreted. FS > 0.9 indicates good fit between true and sample factor structure (Bortz, 1998).
dependent. Missing values were excluded pairwise. The scree plot suggested that eight factors should be extracted. Indeed, this solution was readily interpretable. We also tested solutions with nine, seven, and fewer factors; all were harder to interpret or seemed already to form second-order factors.

The eight-factor solution explained 50.27% of the total variance. A very strong first factor was followed by factors with lower eigenvalues (table 1). The ninth factor accounted for only an additional 2.40% of the total variance.

### 5.3.3 First-Order Factors

Three factors containing subjective descriptions of the own experiences were found (factors 1, 3, and 8 in table 1). Factor 1 was composed of items following the classic description of presence. In fact, the highest-loading item is the common definition of presence: “In the virtual environment I had a sense of being there . . .” (Slater et al., 1994, \( a_{SP} = 0.808 \)). Further items were, for instance, “I felt present in the virtual space” (\( a_{SP} = 0.794 \)) and “I had a sense of acting in the virtual space instead of operating something from the outside” (\( a_{SP} = 0.790 \)). We termed this factor **spatial presence** (SP).

Factor 3 consisted of items describing the attention to real and virtual environment: “I concentrated only on the virtual space” (\( a_{INV} = 0.742 \)), “I was completely captivated by the virtual world” (\( a_{INV} = 0.732 \)), and “I still paid attention to the real environment” (\( a_{INV} = -0.706 \)). Following Witmer and Singer (1998), we called it **involvement** (INV).

Factor 8 also described the subjective experience with items like “How real did the virtual world seem to you?” (Carlin et al., 1997, \( a_{REAL} = -0.677 \)), and “How much did your experience in the virtual environment seem consistent with your real world experience?” (Witmer & Singer, 1994, \( a_{REAL} = -0.591 \)). The items all referred to a comparison between the virtual and the real world, or to a reality judgment. We called this factor **realness** (REAL).

Factors 5, 6, and 7 consisted of items describing the interaction with the environment. Factor 5 evaluated distractions by the interface (interface awareness, IA): “Overall, how much did you focus on using the display and control devices instead of the virtual experience and experimental tasks?”, and “How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?” (ibid., \( a_{IA} = -0.759 \) and \( a_{IA} = -0.625 \), respectively). Factor 6 mainly described how easy the exploration of the VE was: “How closely were you able to examine objects?” (ibid., \( a_{EXPL} = 0.760 \)), and “How well could you examine objects from multiple viewpoints?” (ibid.,

<table>
<thead>
<tr>
<th>Component</th>
<th>Name</th>
<th>Label</th>
<th>Number of Items</th>
<th>Eigenvalue</th>
<th>% of Variance Explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spatial presence</td>
<td>SP</td>
<td>14</td>
<td>14.087</td>
<td>20.717</td>
</tr>
<tr>
<td>2</td>
<td>Quality of immersion</td>
<td>QI</td>
<td>8</td>
<td>4.574</td>
<td>6.726</td>
</tr>
<tr>
<td>3</td>
<td>Involvement</td>
<td>INV</td>
<td>10</td>
<td>3.824</td>
<td>5.624</td>
</tr>
<tr>
<td>4</td>
<td>Drama</td>
<td>DRA</td>
<td>7</td>
<td>3.083</td>
<td>4.533</td>
</tr>
<tr>
<td>5</td>
<td>Interface awareness</td>
<td>IA</td>
<td>7</td>
<td>2.485</td>
<td>3.655</td>
</tr>
<tr>
<td>6</td>
<td>Exploration of VE</td>
<td>EXPL</td>
<td>6</td>
<td>2.262</td>
<td>3.326</td>
</tr>
<tr>
<td>7</td>
<td>Predictability &amp; interaction</td>
<td>PRED</td>
<td>6</td>
<td>1.967</td>
<td>2.893</td>
</tr>
<tr>
<td>8</td>
<td>Realness</td>
<td>REAL</td>
<td>5</td>
<td>1.901</td>
<td>2.795</td>
</tr>
</tbody>
</table>

Table 1. Factor Analysis Study 1. Factors, Numbers of Items, and Explained Variance

---

5. We realize that it would be desirable to provide the full loading structure in a table. Due to space constraints, however, we restrict ourselves to providing loadings on the described factors, and the loading structure for the final measurement model. The complete matrices for the factor analyses can be obtained from the first author.
EXPL = 0.620); hence, we called it the exploration (EXPL) factor. Factor 7, predictability and interaction (PRED) described the interaction with the VE concerning how well its results could be predicted and how broadly it actually took place, with items such as “Were you able to anticipate what would happen next in response to the actions that you performed?” (ibid., $a_{\text{PRED}} = 0.648$), and “How responsive was the environment to actions that you initiated (or performed)?” (ibid., $a_{\text{PRED}} = 0.590$).

The remaining two factors concerned evaluations of the immersion provided by the VE. Factor 2 described the quality of the immersion (QI), e.g. “How much did the auditory aspects of the environment involve you?”, “How well could you identify sounds?”, and “Was the information provided through different senses in the virtual environment . . . consistent?” (ibid., $a_{\text{QI}} = 0.790$, $a_{\text{QI}} = 0.789$, and $a_{\text{QI}} = 0.718$, respectively). Factor 4 evaluated the dramatic involvement (DRA): “Did the virtual world seem to you like a film you were acting in?” ($a_{\text{DRA}} = 0.719$), and “Did you see a plot or a story in the virtual world?” ($a_{\text{DRA}} = 0.668$).

### 5.3.4 Second Order Factors

To interpret the factors, categorizing them on the basis of their correlation in a second-order factor analysis provides additional information. For this purpose, eight scores were created by summing the items belonging (that is, loading highest on this factor and above 0.40) to the respective subscales. Thus, for every participant, we computed eight new scores. These scores were again factorized using principal components analysis and oblique direct Oblimin rotation. The KMO measure had an acceptable value of 0.805. The screen plot suggested a very strong first factor. Because we wanted to test how the factors group together, we report the solutions with two and three factors (explaining 53.82% and 65.60% of variance), although the third factor had an eigenvalue slightly below 1.0. Eigenvalues, explained variances, and loadings on the second-order factors are shown in Table 2.

The three components REAL, SP, and INV together loaded on the first strong second-order factor. When a two-factor solution was forced, they were joined by DRA and QI. In the three-factor solution, these two

### Table 2. Second-Level Factor Analysis, Study 1. Structure Matrices, Eigenvalues and Explained Variance for Two- and Three-Factor Solution

<table>
<thead>
<tr>
<th>First-Order Factors</th>
<th>Three-Factor Solution Component</th>
<th>Two-Factor Solution Component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Realness</td>
<td>0.860</td>
<td>0.700</td>
</tr>
<tr>
<td>Involvement</td>
<td>0.790</td>
<td>-0.546</td>
</tr>
<tr>
<td>Spatial presence</td>
<td>0.787</td>
<td>-0.817</td>
</tr>
<tr>
<td>Interface awareness</td>
<td></td>
<td>0.467</td>
</tr>
<tr>
<td>Predictability &amp; interaction</td>
<td>0.791</td>
<td>-0.791</td>
</tr>
<tr>
<td>Exploration</td>
<td>0.823</td>
<td>0.650</td>
</tr>
<tr>
<td>Drama</td>
<td>0.811</td>
<td>0.622</td>
</tr>
<tr>
<td>Quality of immersion</td>
<td>3.228</td>
<td>1.077</td>
</tr>
<tr>
<td>% of variance explained</td>
<td>40.356</td>
<td>13.468</td>
</tr>
</tbody>
</table>

Loadings below 0.40 were omitted.
variables together formed the third factor. In both solutions, the second factor consisted of IA, PRED, and EXPL.

5.4 Discussion

We asked our participants about their experiences with virtual environments, using previously published and new items, and then factorized these results. The results of the factor analysis show that the items used in the survey split into distinct factors, describing either presence, immersion, or interaction. This interpretation is supported by the second-order factor analysis, which combined the factors in this manner.

5.4.1 Presence Factors. Among the eight extracted factors, three combine items describing the subjective interpretation of the own experiences. The SP factor is a clear manifestation of the accepted definition of presence as the sense of being there, underlined by the fact that the commonly used phrase “sense of being there” is the highest-loading item on this factor. Other items on this factor emphasize the importance of actions in the VE: for instance, “I had a sense of acting in the virtual space, rather than operating something from outside.” The INV factor combines items describing awareness and attention processes. It seems that we have here a manifestation of the attention component of the presence experience, as proposed by Witmer and Singer (1998) and as predicted by our cognitive analysis. These two factors also load together on the first second-order factor, indicating that this is a general presence factor. Additionally, we find realness loading on this factor. Its items tap judgements of the VE concerning its realness or comparability to reality. We return to this factor in the discussion of study 2.

5.4.2 Immersion and Interaction Factors. In contrast to the presence factors, all other factors tap descriptions of the stimuli offered by VE and the interface, and the interaction with them. It seems that our factor structure matches previous categorizations very well.

Overall, the factor QI fits the sensory factors presented by Witmer and Singer (1998). Three of the four elements of sensory factors cited by them can be found in the QI factor: environmental richness (Sheridan, 1992), multimodal presentation (Held & Durlach, 1992), and consistency of multimodal information (ibid.). In a more general sense, the QI can be related to Steuer’s (1992) vividness construct: “the ability of a technology to produce a sensorially rich mediated environment” (p. 80).

Items in the IA factor seem to match Held and Durlach’s (1992) argument that unnatural and distracting interfaces interfere with the emergence of presence. When the interface is easy to use, the user can concentrate on the activities, quickly adapt to the VE, and use it proficiently (Witmer & Singer, 1998).

EXPL seems to match what Witmer and Singer call active search: “an environment should enhance presence when it permits observers to control the relation of their sensors to the environment” (Sheridan, 1992, p. 230). It constitutes close and complete exploration of objects and the environment from multiple viewpoints. It is also important that these interactions seem natural to the user.

Our PRED factor matches what Witmer and Singer call anticipation. Anticipation is made possible by a mental model of the dynamics inside the virtual environment, a point Slater et al. (1994) mention as one immersion factor: “the connection between a participant’s actions and effects should be simple enough for the participant to model over time” (p. 131).

Together, EXPL and PRED are marked by the impression that actions can be performed successfully: it is important that actions can be performed in the first place (basic interactivity), that the actions have an impact on the VE, and that this impact is the one that is desired and predicted. Interestingly, a more general knowledge of the VE concerning its spatial and dynamic properties is part of that.

The DRA factor is related to the description of plot impacts given by Slater and Wilbur (1997): “[plot] is the extent to which the VE in a particular context presents a story-line that is self-contained, has its own dynamic, and presents an alternate unfolding sequence of events” (p. 605). The factor is constituted by the perception of a dramatic storyline comparable to a movie.
or a book, in which the user can participate. The perception of dramatic moments, unexpected highlights, and exciting events is also important. However, a stronger item asking “Did you have a sense of being able to change the plot?”, breaks out of this factor and loads on the PRED factor. Thus, it seems that the DRA factor relates more to the “passive” perception of the plot offered by the VE.

We collected these data mainly on the Internet, with a self-selecting sample. The resulting heterogeneity concerning applications and technologies broadens the focus of our analysis. Unfortunately, our approach also resulted in three major asymmetries in the sample structure. First, our participants were of course established Web users. Second, the overwhelming majority of the participants were male. Third, they mainly described games played on a desktop computer with a monoscopic display. Could this have distorted our results? We do not think so. First, there is no reason to assume that Web users experience VEs different from non-Web users. Second, from a cognitive perspective, there is no reason to assume qualitatively different spatial-constructive processes in men and women. Furthermore, Lessiter, Freeman, Keogh, and Davidoff (2000) present results that are similar to our findings, although they were based on a more heterogeneous sample. With respect to the dominance of games, it is ensuring that previous conceptualizations fit our results very well. Consider the spatial presence factor: it is composed of items that fit very clearly the classic definition of presence. In the same manner, involvement and the immersion factors fit previous theorizing and results collected with more-different samples and more-advanced technologies. However, it may be possible that the realness factor emerges especially when dealing with the games used by our participants, because they often deliberately negate the constraints of reality. But, even if our results were slightly influenced by the types of applications used, one should not forget that, at the present time, 3-D games are probably the most common VEs, and thus an important topic themselves.

6 Study 2

Factor analysis is rather exploratory than hypothesis testing, and replication of factor structures is crucial for obtaining reliable results. Therefore, study 2 was primarily designed to replicate central aspects of study 1. The presence and interaction factors were the focus of study 2, omitting the immersion factors found in study 1.

6.1 Materials and Procedure

Not all items from study 1 loading on the identified factors were included in the survey. First of all, only items belonging to the three presence factors, and items belonging to EXPL and PRED were included, because our theoretical focus was especially on presence and interaction. Second, based on the results of study 1, preliminary scales were formed out of items loading on the same factor, excluding some of their items. The item selection balanced two concerns: although the content of each scale should vary as much as possible and include different facets, especially highly loading items were chosen. The numbers of items were ten items for SP, ten for INV, five for REAL, six for EXPL, and six for PRED.

Additionally, addressing shortcomings of study 1, study 2 explicitly asked whether a first-person perspective or a third-person perspective was used, and when exactly the described interaction episode took place. To add an incentive for completing the questionnaire, one randomly chosen participant won a graphics board in a lottery among all participants.

6.2 Participants

Two different versions of the questionnaires were offered. Depending on whether the environment contained narrative content, additional items on the perception of the narration were added, with the nonnarrative version being a subset of the narrative version. We will report data for the nonnarrative items only, which were collected with both versions. For these variables, data from 296 participants were collected.
Only 11.2% of the participants were female, and 88.2% male. The mean age was 24.7 years (SD = 6.2), ranging from 13 to 50 years. The majority (n = 227) experienced the VE on a monoscopic monitor, and 33 used HMDs. When asked for their dominant perspective on the VE, 190 indicated that they had a first-person perspective, and 87 answered that they had a third-person perspective (missing data for 19 cases). The main type of application were 3-D games (n = 252), but some (n = 6) worked with visualizations (missing data for 38 cases). We also asked when the interaction episode described by the participants took place: 53 described an episode that happened on the same day, 162 described an episode that took place some days earlier, and for 81 the episode took place more than a week before completing the questionnaire.

### 6.3 Results

**6.3.1 Sample Adequacy and Factorization.**

The various measures for sampling adequacy again indicated good preconditions for factor analysis (296 subjects, ratio of subjects to variables 296:37 (8:1), ratio of subjects to factors 296:5 (59.2), FS index = 0.927, KMO = 0.923). The sample was analyzed following the same procedures as in study 1. The scree plot suggested the extraction of five factors, which was consistent with the results from study 1. This solution explained 53.50% of the total variance. Two additional factors had eigenvalues higher than 1. The sixth factor would have accounted for an additional 3.0% of the total variance. See table 3.

<table>
<thead>
<tr>
<th>Component</th>
<th>Name</th>
<th>Label</th>
<th>Number of Items</th>
<th>Initial Eigenvalue</th>
<th>% of Variance Explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spatial presence</td>
<td>SP</td>
<td>14</td>
<td>11.767</td>
<td>31.804</td>
</tr>
<tr>
<td>2</td>
<td>Exploration of VE</td>
<td>EXPL</td>
<td>8</td>
<td>2.682</td>
<td>7.248</td>
</tr>
<tr>
<td>3</td>
<td>Realness</td>
<td>REAL</td>
<td>3</td>
<td>2.070</td>
<td>5.594</td>
</tr>
<tr>
<td>4</td>
<td>Predictability &amp; interaction</td>
<td>PRED</td>
<td>2</td>
<td>1.717</td>
<td>4.640</td>
</tr>
<tr>
<td>5</td>
<td>Involvement</td>
<td>INV</td>
<td>10</td>
<td>1.559</td>
<td>4.214</td>
</tr>
</tbody>
</table>

*Table 3. Factor Analysis Study 2. Factors, Numbers of Items, and Explained Variance*

**6.3.2 Factors.** The factor structure largely confirmed that of study 1: the same five factors as in study 1 could be identified. The comparison of the factor structures between study 1 and study 2 revealed two points. First, the presence factor structure was relatively stable. Both SP and INV kept all their items. All top eight items of SP came from the study 1 SP component. However, two items from the study 1 REAL factor loaded highest on SP. Thus, REAL was left with only three items. These three items all asked explicitly for a comparison between real and virtual environment.

Second, the structure for the interaction items changed in some respects. One item from each, PRED and EXPL, moved to SP (“I knew what the virtual world behind me looked like” and “How natural did your interactions with the environment seem?”). Also, PRED lost three more items, namely its interaction part, to EXPL (for example, “How much were you able to control events?”). Only the two central “prediction” items were left in PRED.

### 6.4 Discussion

The factor structure extracted in study 2 confirms the central results of study 1 and clarifies others. First, the factors SP and INV were almost identical to those from study 1. Both their content and the division between these factors could be replicated. This is important because it confirms our basic prediction that an attentional and a spatial constructive component are experienced due to the cognitive processes leading to presence.
In study 2, we again found a factor that combined items asking for the realness of the VE. This factor was a surprise to us; furthermore, no comparable cluster was found by Witmer and Singer (1998). In study 1, all questions loading on REAL except one asked for a comparison between real environment(s) and the VE. 6 In study 2, the factor is altogether constrained to this meaning. This seems to clarify its meaning. Although it is not clear whether it is a judgment elicited by our questions or a true component of the presence experience, the idea that attribution of reality or realness is a part of the sense of presence has been advocated earlier. Slater et al. (1994) wrote that “the extent to which, while immersed in the VE, it becomes more ‘real or present’ than everyday reality” might be an indicator for presence (p. 132). Similarly, Sheridan (1996) proposed to use “verbal descriptions of the experience or the degree of realism experienced” (p. 242). But it is, in general, acknowledged that nobody would really mistake a VE for reality (Steuer, 1992).

The facts that REAL is a separate factor and that it loads with INV and SP on a general presence factor show that it is both different from what is commonly understood as presence and closely related to it. It may, in fact, be the “ingenuous realism” of the users (Mantovani & Riva, 1999) at work.

Two items from factors formerly identified as interaction factors (PRED and EXPL) now load highest on SP. This points out that presence experiences and evaluations of interaction and immersion may be inseparable in some items due to their wording. In exploratory factor analyses, such a confusion results in double loadings. To get purer factors, confirmatory factor analysis as used below is more useful and exact, because double loadings are strictly controlled.

The sample of study 2 was very similar to that of study 1, and thus it is again asymmetric concerning age and types of applications. Two-thirds of the participants had a first-person perspective on the VE, and one-third steered a vehicle or a character through the VE. We do not think that this renders the data less valid: everyone who had played a racing game will agree that this can produce high degrees of presence, and users playing third-person-perspective games often identify themselves with the main character and mimic his or her movements.

Also, the data indicate that, although 72.6% of the participants described interactions that happened the same day or only some days ago, 27.3% described interactions that took place more than a week before the survey. Although we see no theoretical reason why a longer time between experience and report should distort the data in the direction of our findings, we conducted additional analyses that did not show important differences between these categories of users.

7 Confirmatory Factor Analysis of a Three-Component Presence Scale

For the presence scales of study 2, we already selected items from the presence factors identified in study 1. The scales formed in this step, however, were still fairly large. To reach more-parsimonious scales and to identify items that show fewer double loadings, we computed a confirmatory factor analysis (CFA). The mathematical basis of CFA is closely related to exploratory factor analysis: a factor is considered as a latent variable, which is estimated from the observed covariance matrix. In contrast to exploratory factor analysis, however, CFA offers an assessment of the fit between model and observed data. We used CFA here in a multi-step procedure to reach a fitting model. It is, however, important to note that a fitting model can be adapted to nearly every data set. The crucial point is that the model also must fit a second independent data set. We therefore developed a model on the basis of the study 1 data and then tested its fit to the study 2 data.

The first step in the model building was simply to enter the factor structure observed in study 1 into a CFA model. We included all items that were selected after study 1: that is, ten items for SP, ten for INV, and five for REAL. Instead of simply letting the three factors

---

6. The remaining question uses the German term *wirklich*, which is hard to translate. Its meaning is close to actual, genuine, and real, but it somehow already acknowledges that this reality may be psychologically construed.
(latent variables) covary, we introduced a second-level factor (PRES) reminiscent of the second-level factor analysis presented above. The model is illustrated in figure 1 on the left side. Note that each factor (the latent variables, drawn as circles) loads on only its own items, not on items belonging to other factors. This is the difference to the exploratory factor analyses, in which double loadings were not controllable. Because modification indices cannot be computed when values are missing, we imputed the missing values by simply replacing them with the mean of the respective variable.

Due to double loadings, this model—although the outcome of a factor analysis—did not fit the data, as indicated by two fit statistics: the implied covariance matrix differed significantly from the observed matrix, $\chi^2(272, N = 246) = 751.592, p < 0.001$, RMSEA = 0.085. This is due to double loadings.

The model was then adapted following two rules:

1. Items that loaded on all three factors and the general factor PRES (as indicated by modification indices above 5 for these paths) were drawn out and received only a path from the general factor PRES.

2. Items that had loadings from other factors than their own (modification indices above 5) were deleted, as long as rule 1 did not apply.

In short, our procedure eliminated the items that were likely to measure two types of presence at once, but not all three at once. These rules were applied in three iterations. Rule 1 did apply to two variables, but one had to be deleted due to unresolvable double loadings with INV (application of rule 2). In the last step, one more variable was deleted because it had strong modification indices on regressions to other variables. The resulting model had an excellent fit, $\chi^2(62, N = 246) = 68.628, p = 0.263$, RMSEA = 0.021. The model features one item loading on the general factor. This item is, not surprisingly, “I had a sense of being there . . .”, which was formerly the highest loading item on the SP factor. Five items remain on SP, four items on INV, and three items on REAL (figure 1, right model). The items and their loadings can be seen in table 4.

As a second independent data set, study 2 data were then entered into the model, with missing values again replaced by means. The model fitted these data as well, $\chi^2(62, N = 296) = 89.840, p = 0.012$, RMSEA = 0.073.

---

7. Root mean square error (RMSEA) is a recommended test statistic for confirmatory factor analyses or structural equation models in general. It has a range from 0 to 1. Model fit is acceptable with RMSEAs below 0.8 and good with RMSEAs below 0.5 (Browne & Cudeck, 1993).

8. With this model, a parsimonious three-component presence scale is reached. For application purposes, however, we would recommend the addition of one item to the REAL subscale.
Table 4 Standardized Loadings in the Three-Component Presence Model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Criterium</th>
<th>Abbreviated item</th>
<th>Standardized Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Study 1</td>
</tr>
<tr>
<td>PRES</td>
<td>SP</td>
<td>spatial presence (latent variable)</td>
<td>0.852</td>
</tr>
<tr>
<td>PRES</td>
<td>INV</td>
<td>involvement (latent variable)</td>
<td>0.474</td>
</tr>
<tr>
<td>PRES</td>
<td>REAL</td>
<td>realness (latent variable)</td>
<td>0.672</td>
</tr>
<tr>
<td>PRES</td>
<td>G1</td>
<td>sense of being in a place</td>
<td>0.876</td>
</tr>
<tr>
<td>SP</td>
<td>SP1</td>
<td>sense of VE continuing behind me</td>
<td>0.491</td>
</tr>
<tr>
<td>SP</td>
<td>SP2</td>
<td>sense of seeing only pictures</td>
<td>-0.566</td>
</tr>
<tr>
<td>SP</td>
<td>SP3</td>
<td>sense of being in the virtual space</td>
<td>0.663</td>
</tr>
<tr>
<td>SP</td>
<td>SP4</td>
<td>sense of acting in the VE</td>
<td>0.845</td>
</tr>
<tr>
<td>SP</td>
<td>SP5</td>
<td>felt present in the VE</td>
<td>0.808</td>
</tr>
<tr>
<td>INV</td>
<td>INV1</td>
<td>awareness of real world stimuli</td>
<td>-0.685</td>
</tr>
<tr>
<td>INV</td>
<td>INV2</td>
<td>awareness of real environment</td>
<td>-0.647</td>
</tr>
<tr>
<td>INV</td>
<td>INV3</td>
<td>attention to the real environment</td>
<td>-0.671</td>
</tr>
<tr>
<td>INV</td>
<td>INV4</td>
<td>captivated by the VE</td>
<td>0.624</td>
</tr>
<tr>
<td>REAL</td>
<td>REAL1</td>
<td>How real seemed VE in comparison with the real world?</td>
<td>0.781</td>
</tr>
<tr>
<td>REAL</td>
<td>REAL2</td>
<td>consistency of experiencing the VE and a real environment</td>
<td>0.553</td>
</tr>
<tr>
<td>REAL</td>
<td>REAL3</td>
<td>How real seemed VE in comparison with an imagined world?</td>
<td>0.572</td>
</tr>
</tbody>
</table>

Items INV2, INV3, and REAL2 are actually anchored reverse, but their loadings were multiplied by -1 for ease of interpretation. Loadings of PRES on G1, SP on SP5, INV on INV4, and REAL on REAL3 were fixed to achieve identifiability, which prevents computation of significances for these variables. All other loadings are highly significant at \( p < 0.001 \). Item G1 was taken from Slater and Usoh (1994), items INV1 and REAL2 from Witmer and Singer (1994), item REAL1 from Hendrix (1994), and item REAL3 from Carlin et al. (1997).

Although the chi-square test was significant, it is known that this test gets exceedingly imprecise with larger samples (Browne & Cudeck, 1993). The more reliable and preferable RMSEA indicated a good fit of the model. Table 4 shows the items’ standardized loadings on the factors, interpretable similar to correlations.10

9. The structural equation modeling software AMOS 4 also allows the maximum likelihood estimation of missing values, which is more exact than simply replacing missing values with means, but excludes estimation of modification indices. Estimating the model fit with this option, it does not change substantially; \( \chi^2(62, N = 296) = 90.822, p = 0.010 \), RMSEA = 0.040.

10. The item labels in table 4 point to the item names in the documents at [http://www.igroup.org/pq/ipq/](http://www.igroup.org/pq/ipq/). From this site, all items of the final model in German and English, as well as the respective data from study 1 and 2, can be downloaded.

8 General Discussion

What are the insights from our factor analyses? We presented factor analyses for two studies of presence and immersion experiences. Concerning the three goals outlined earlier, we can summarize that
our interpretation of the results suggests that the items used in the surveys split into distinct factors, describing either subjective presence experiences, evaluations of the immersive technology, or evaluations of the interaction. This interpretation is supported by the second-order factor analysis, which combined the factors in this manner. Our central prediction that presence experiences involve two distinct components—namely spatial-constructive and attention facets—was confirmed. Additionally, we found a third subjective component: judgments of realness. Concerning evaluations of immersion and interaction, we found five factors, and the structure of these factors supports previous categorizations of factors influencing presence.

8.1 Comparison to Other Studies

These results both support and extend the arguments of Witmer and Singer (1998). Because our results differ from their cluster analysis, we have to point out that most of the items constituting the presence factors were not part of Witmer and Singer’s questionnaire. Out of the 24 items constituting the SP and INV factors in study 1, only four come from Witmer and Singer (1994). This also supports the critique of Slater (1999), who argues that, although the Witmer and Singer questionnaire contains subjective evaluations of the immersion provided by the system, it does not in fact measure presence experiences, but subjective evaluations of its contributing factors. Results of a path analysis between the immersion evaluations found in study 1 and presence experiences (Schubert et al., 1999) show that, although these components are separable, presence is nonetheless predicted by immersion and interaction evaluations. This may account for the correlations reported by Witmer and Singer (1998).

Our results are supported by a recent factor analysis presented by Lessiter et al. (2000). In an analysis of 63 items, they found four factors: physical space (“I had a sense of being in the scenes displayed”), engagement (“I felt involved [in the displayed environment]”), naturalness (“The content seemed believable to me”), and negative effects (“I felt dizzy”). As the authors also point out, the first three factors seem very similar to our factors SP, INV, and REAL. The last factor includes symptoms of simulator sickness, which were not included in the analyzed items of study 1 and 2. Most interestingly, of the 604 participants in this study, only 49 played on a videogame console before answering the questionnaire, whereas the remaining participants experienced IMAX cinema or other forms of noninteractive content. Moreover, in contrast to our sample, participants were more equally distributed across ages and gender. Therefore, we take the results from Lessiter et al. as strong indication that the results we found hold across different media and participants, which supports the notion that they tap underlying cognitive processes instead of mere media characteristics.

8.2 New Types of Hypotheses

Acknowledging that presence is a multidimensional construct prompts us to ask new questions in presence research. The prototypical hypothesis in today’s presence research regresses a unitary presence measure on one or several characteristics of the VE, and many of these hypotheses are developed “mainly on intuitive grounds” (Welch et al., 1996, p. 265).

Two other types of hypotheses are possible: one could ask which characteristic of the VE should determine which presence components, and which not. In two recent studies, we have shown that providing both real and illusory possibilities to interact with a virtual environment through bodily movement enhances spatial presence, but only marginally involvement and realness (Schubert, Regenbrecht, & Friedmann, 2000). The second new type of question could ask which characteristic of the VE should influence which mediating cognitive process, which in turn determines different presence components. Our results show that the two processes predicted from the embodied cognition framework surface in subjective experiences. Hypotheses that combine investigation of how users perceive, conceptualize, and experience VEs may be the venue to a cognitive theory of presence in virtual environments.
Acknowledgments

We wish to thank the members of the igroup for their help in conducting this research, VRT GmbH for its financial support, and Sven Waldzus and three anonymous reviewers for helpful comments on earlier drafts of this article.

References


