Jane Lessiter
j.lessiter@gold.ac.uk

Jonathan Freeman
j.freeman@gold.ac.uk

Edmund Keogh
Jules Davidoff
Department of Psychology
Goldsmiths College
University of London
New Cross, London SE14 6NW
United Kingdom

A Cross-Media Presence Questionnaire:

The ITC-Sense of Presence Inventory

Abstract

The presence research community would benefit from a reliable and valid cross-media presence measure that allows results from different laboratories to be compared and a more comprehensive knowledge base to be developed. The ITC-Sense of Presence Inventory (ITC-SOPI) is a new state questionnaire measure whose development has been informed by previous research on the determinants of presence and current self-report measures. It focuses on users' experiences of media, with no reference to objective system parameters. More than 600 people completed the ITC-SOPI following an experience with one of a range of noninteractive and interactive media. Exploratory analysis (principal axis factoring) revealed four factors: Sense of Physical Space, Engagement, Ecological Validity, and Negative Effects. Relations between the factors and the consistency of the factor structure with others reported in the literature are discussed. Preliminary analyses described here demonstrate that the ITC-SOPI is reliable and valid, but more rigorous testing of its psychometric properties and applicability to interactive virtual environments is required. Subject to satisfactory confirmatory analyses, the ITC-SOPI will offer researchers using a range of media systems a tool with which to measure four facets of a media experience that are putatively related to presence.

1 Introduction

1.1 Defining Presence

Presence has been used as a global experiential quality metric to evaluate, develop, and optimize both advanced broadcast and virtual environment (VE) media systems (Freeman & Avons, 2000; Freeman, Avons, Pearson, & IJsselsteijn, 1999; Slater, Usoh, & Steed, 1994). *Presence* is generally defined as a user's subjective sensation of "being there" in a scene depicted by a medium (Barfield, Zeltzer, Sheridan, & Slater, 1995). It has also been defined as "a perceptual illusion of non-mediation" (Lombard & Ditton, 1997), a definition that is consistent with the former one as it implies that a user incorrectly perceives a mediated scene to be unmediated. Further, Slater and Usoh (1994) described presence as "the (suspension of dis-) belief" of being located in a world other than the physical one.

In addition to these specific definitions, presence has been variously described as a "mental manifestation" (Sheridan, 1992a), a (general) "existential phenomenon" (Draper, Kaber, & Usher, 1998), and a "cognitive state" consistent with a sense of "being there" in an environment, a state that results

Presence, *Vol. 10, No. 3, June 2001*, 282–297

from attending to and evaluating incoming sensory information (Barfield et al., 1995).

Others have used the terms *telepresence* (Barfield et al., 1995; Sheridan, 1992a; Welch, Blackmon, Liu, Mellers, & Stark, 1996) and *virtual presence* (Barfield et al., 1995) to refer to presence in relation to specific technologies. Draper et al. (1998) define telepresence as "the perception of *presence* within a physically remote or simulated site," which suggests, de facto, that presence is a valid construct in relation to experience of the real (physical) world. This is a contentious issue, a full discussion of which is beyond the scope of this paper. However, it is our view that presence is a more useful concept when it is limited to the study of users' experiences of mediated presentations. Real-world experience can be adequately described in terms of more traditional psychological constructs: such as attention, involvement, and arousal, to name but a few. There seems little to gain from describing people's everyday experience in terms of presence. In the physical world, "being there" is an invariant, and any variations in sensations of "being there" are rarely contemplated. Consistent with the definitions given above, the reason that presence is relevant to understanding users' experiences of media is that an illusion is generated whereby a user senses that she/he is located somewhere other than her/his physical environment. This illusion can be entertaining (Lodge, 1999), but in addition it might affect users' performance within mediated environments (Welch, 1999) and support effective therapeutic applications (North, North, & Coble, 1997). It remains though, at all times, an illusion. At the very least, though, real-world experience is useful to presence research insofar as it serves as a benchmark, or standard, against which to subjectively judge levels of presence in mediated environments.

Whether there are qualitative differences in the sensations supported by different media systems is an open question. However, in a focus group study designed to investigate viewers' experiences of stereoscopic television, Freeman and Avons (2000) found that observers used terms commonly associated with virtual environment display systems, such as "being there" (cf. Usoh et al., 1999). This suggests that presence may be evoked by a range of media to varying degrees. As presence is multiply determined, it is likely that tradeoffs exist between the different determinants. For instance, displays that support a high degree of photorealism (such as cinema or high-definition TV) may compensate for an absence of control and manipulation input devices (henceforth, referred to as *interactive* displays) typically supported in VEs.

Through a variety of self-report measures, evidence in support of a range of determinants of presence has been generated. To more easily compare the relative impact on presence of its determinants, a general (cross-media), valid, and reliable measure of presence is desirable. This paper describes the development of the ITC-Sense of Presence Inventory (ITC-SOPI), which we propose as one such measure.

1.2 Determinants of Presence

It is beyond the scope of this paper to provide a thorough review of all the research that has been completed into the determinants of presence. For excellent reviews, the reader is referred to Barfield et al. (1995) and Draper et al. (1998). Two explanatory approaches to presence have been used to date: those based on properties of the technology, and those that compare and contrast other psychological experiences with presence (Draper et al., 1998). Partly as a result of these divergent approaches, the literature has on occasion confused determinants and correlates of presence. An additional explanation for this confusion is the genuine debate as to whether some factors (for example, interest and engagement) are determinants or correlates of presence. These issues will not be resolved here. For the purposes of this paper, we will identify content areas relevant to the development of the ITC-Sense of Presence Inventory.

Two general categories of variables can determine a user's presence: media characteristics and user characteristics. This differentiation shares parallels with that of Slater and colleagues (Slater, Steed, McCarthy, & Maringelli, 1998; Slater & Usoh, 1994; Slater & Wilbur, 1997) who described "external" (objective) and "internal" (subjective) determinants of presence. The media characteristics category can be subdivided into aspects of media form and media content (for example, IJsselsteijn, de Ridder, Freeman, & Avons, 2000). It is likely that these categories will interact in determining presence.

Media form refers to physical, objective properties of a display medium. Sheridan (1992a) proposed three major elements of media form that may determine presence: the extent of sensory information presented, the degree of control a participant has over positioning his/her sensors within the environment (such as a turn of the head to see or hear more of the environment), and a user's ability to modify aspects of the environment. An ideal teleoperation system could support each of these elements sufficiently to make the mediated presentation indistinguishable from the physical world, and Sheridan's theory would predict that such a system would elicit high presence. By this definition, media form covers the means by which an image is represented (for example, photorealistic video or animated computer graphics).

We use the media content category to refer to the overall theme, narrative, or story depicted via a display system. In experiments conducted to date, this has ranged from a virtual cliff scenario (Usoh et al., 1999) to a fast-paced rally car drive (Freeman, Avons, Meddis, Pearson, & IJsselsteijn, 2000; Freeman & Avons, 2000). Clearly, some aspects of a narrative might change as a result of variation in media form variables. For example, in an interactive VE, a user can select the direction in which to navigate; hence, the order of events within the VE may vary. However, the overall theme—what we refer to as *media content*—remains unaffected.

Social elements of a displayed environment, such as acknowledgement of the user through the reactions of other actors, virtual or real (Heeter, 1992), also contribute to determining presence (sometimes referred to as *social presence*). The reader is referred to Lombard and Ditton (1997) for a comprehensive review of research into social aspects of presence. Different display configurations are likely to be required to optimize media systems for social presence rather than (spatial) presence. Some media systems, such as the immersive teleconferencing system being developed by the European Commission's IST VIRTUE project (2000), aim to support both high presence and high social presence.

There is some evidence that user characteristics, such as a user's perceptual, cognitive and motor abilities, and personality traits (such as a willingness to suspend disbelief (Slater & Usoh, 1994)), can be important in determining presence (Witmer & Singer, 1998). Relevant individual characteristics are likely to vary with the age and possibly with the sex of the user. Further, as presence is a transient experiential state, like mood (Sheridan, 1992a, 1992b), it is susceptible to variation within the same person, given the same physical conditions on two separate occasions.

Slater and colleagues have reported preliminary investigations into the relationships between individual differences and presence (Slater et al., 1998; Slater & Usoh, 1994; Slater & Wilbur, 1997). Using both subjective and objective methodologies, Slater and colleagues have demonstrated that users whose primary representation systems are visual are more likely to experience presence in a visual virtual environment than users whose primary representation systems are auditory or kinaesthetic.

Witmer and Singer (1998) made an important contribution to the measurement of individual differences in presence. Their Immersive Tendencies Questionnaire contains 29 items of which 16 relate to one of three clusters (involvement, focus, and games) that tap both state (for example, "How mentally alert do you feel at the present time?") and trait ("Have you ever remained apprehensive or fearful long after watching a scary movie?") individual differences that putatively moderate the tendency to feel present.

Attention and involvement are user responses associated with presence and are likely to be affected by all three of the categories of determinants previously outlined: media form, media content, and user characteristics. Involvement and attention are particularly important in the measurement of presence (Barfield & Weghorst, 1993; Witmer & Singer, 1998) as participants exposed to a media system and unfamiliar with the presence concept are likely to use these terms to describe their experience. For instance, Freeman and Avons (2000) found that participants described their involvement in a 3-D video presentation as "unavoidable". The term *flow* has been used to describe a state of augmented concentration, in which the user is unaware of external distractors, the placement of self in the real world, and even real time (Draper et al., 1998; Fontaine, 1992).

To summarize, several empirical studies have demonstrated that presence tends to increase as the fidelity of a reproduction or simulation of the physical world increases.

High-fidelity representations of environments are perceived as being more natural and real—having depth, space, and continuity—and, consequently, as less distracting. The naturalness of a visual representation (IJsselsteijn, de Ridder, Hamberg, Bouwhuis, & Freeman, 1998) and of a user's interaction with a media system (Hendrix & Barfield, 1996a) have both been shown to correlate with presence. When more attention is allocated to the mediated environment than to the environment in which an individual is physically located, the increased mediated sensory input and decreased nonmediated sensory input give rise to sensations of being located within the mediated environment (Kim & Biocca, 1997). Presence involves the user feeling as though they are spatially located within an environment portrayed by a display system, and it is characterized by a sensation of a strong perception-action link between the display and the user. Which of the categories of determinants of presence previously reviewed is most important is a question that cannot be answered at present due to the range of different presence measures employed.

1.3 Subjective Measures of Presence

Sheridan (1992a) reasons that, because presence is a "mental manifestation," its fundamental measurement is self-report. Presence has been measured using simple rating scales (Hendrix & Barfield, 1996b; Slater & Usoh, 1994; Slater et al., 1994; Welch et al., 1996) that typically relate to feeling: physically located in a mediated space, that the mediated environment is as real as the real world, and that the mediated place had been "visited." More structured approaches to the measurement of presence that aim to elucidate the presence construct (and to provide a more reliable and valid index of the dimension(s) related to presence) have also been followed. Table 1 summarizes a range of existing subjective presence measures.

An ideal presence questionnaire that could be used to evaluate a range of media systems must satisfy a number of important criteria. First, an understanding of presence should not be assumed by directly asking respondents how present they feel. Presence is a relatively unfamiliar construct to most nonexperts (Freeman & Avons, 2000; Freeman et al., 2000). Not understanding the construct for which they are providing ratings would reduce participants' ability to meaningfully distinguish points along a presence rating scale (Freeman, 2000).

Second, questions should avoid addressing two issues in one question; this renders the response options confusing and meaningless. Care should be taken in phrasing questions that are answered on a Likert-type scale to ensure that only one (continuous) construct is being measured (for example, "How much did you feel as though the mediated environment was a place that you visited?").

Third, the response options used should, ideally, be consistent across items (within individual questionnaires). This renders the questionnaire more user-friendly and reduces its completion time.

Fourth, presence is likely to be a multidimensional construct (Barfield et al., 1995, Witmer & Singer, 1998; Schubert, Friedmann, & Regenbrecht, 1999). An ideal presence questionnaire should take account of this possibility and tap a range of characteristics that are putatively related to presence. Indeed, unidimensional presence ratings have been shown to be potentially unstable in that they can be affected by prior experience (Freeman et al., 1999). When a group of observers was asked to first rate a training stimulus for how interesting it was, subsequent presence ratings for a test stimulus were less sensitive to variations in viewing condition (the presence or absence of stereoscopic depth cues) than were those of a group asked to rate the same training stimulus for 3D-ness (Freeman et al., 1999). This result was taken to indicate that both depth within the stimulus and interest in the content contribute to pres-

Table 1. *Summary of Subjective Measures of Presence*

	N. items	N. participants	Sample	Analysis	Dimensions	Labels
Slater et al. (1994)	3	N/A	N/A	N/A	N/A	Presence
Slater & Usoh (1994)	1	N/A	N/A	N/A	N/A	Presence
Hendrix & Barfield (1996b)	3	N/A	N/A	N/A	N/A	Presence
Welch et al. (1996)	2	N/A	N/A	N/A	N/A	Presence
Kim & Biocca (1997)	8	96	Users exposed to video stimuli presented on either 9, 20, or 32 in. TV	Factor analysis	2	Arrival Departure
Witmer & Singer (1998)	32	152	VE users	Cluster analysis	3	Involved/Control Natural Interface Quality
Barfield, Baird, & Bjorneseth (1998)	18	Not stated	Not stated	Factor analysis and cluster analysis	Not stated	(subset relate to Presence)
Schubert, Friedmann, & Regenbrecht (1999)	75	246	90% male; mostly 3-D computer games players	Principal components analysis (oblique)	8	Spatial Presence* Quality of Immersion Involvement* Drama Interface Awareness Exploration of VEs Predictability and Interaction Realness (believability)*

* Related to "presence" factor following second-order analysis.

ence, but that the importance of each could be varied by participants' experience immediately prior to providing presence ratings. One potential explanation of the problems of stability and bias associated with simple presence rating scales is that they treat presence as unidimensional when it is in fact multidimensional. Thus, a measure that takes account of the potential multidimensional structure of presence may prove to be more robust.

Fifth, qualitative reports from experimental participants suggest that there are distinct similarities in people's experiences of, and sensations associated with, "being there" elicited by different types of mediated environment (Freeman & Avons, 2000; c.f. Usoh et al., 1999). Hence, a general presence measure that focuses entirely on users' experiences with media presentations should be possible. Thus, to ensure generalizability to other media systems and to avoid confounding separate user and system contributions to presence (Slater, 1999), the questions should not make reference to specific media system and content properties. For example, references to input and interaction devices mean a questionnaire cannot be applied to the evaluation of presence through noninteractive media. Furthermore, depending on their phrasing, questions relating to specific technical devices may not measure presence at all (Slater, 1999). The sense of interacting with a mediated environment can be probed without making direct reference to physical properties of the system. Because a user's perception of, and ease of interaction with, an interface are important issues to consider when designing media systems, there is a good rationale for measuring them with supplementary questions/questionnaires.

Sixth and relatedly, a general presence measure should be piloted on participants of a range of media systems displaying a range of contents, that is, systems with variation in media form and media content.

Finally, standard data-reduction techniques, such as factor analysis, used in the development of questionnaires assume a high respondent-to-item ratio (Kline, 1993). Violation of this criterion reduces the reliability of the correlation coefficients that are used to compute the factors (Tabachnick & Fidell, 1989). Comrey (1973 in Tabachnick & Fidell) suggests that a sample size of 50 is very poor, 100 is poor, 200 is fair, 300 is good, 500 is very good, and 1,000 is excellent. However, as a general index, Tabachnick and Fidell suggest that it is "comforting to have at least five cases for each observed variable" (p. 603).

Our research group has developed a new cross-media presence questionnaire to measure media users' state of presence, which aims to have broad application and adheres to the criteria above. In section 2, we describe the development of the ITC-Sense of Presence Inventory (ITC-SOPI). We then present the results of an exploratory factor analysis, and compare these results to those of other studies that have investigated the structure and measurement of presence.

2 Method

2.1 ITC-SOPI Development

Sixty-three items were initially generated that tapped possible manifestations of different content areas deemed relevant to presence on the basis of theoretical and empirical papers: sense of space, involvement, attention, distraction, control and manipulation (autonomy), realness, naturalness, perception of time, awareness of behavioral responses, a sense of social interaction (parasocial and copresence), personal relevance, arousal, and negative effects.

Any questionnaire should be as respondent-friendly as possible and have face validity; that is, it should look like it is measuring what it purports to measure (Oppenheim, 1992; Rust & Golombok, 1989). A five-point Likert scale (1 = strongly disagree; 5 = strongly agree) was chosen as the response option for all items. This consistency makes it easier for respondents to complete the questionnaire and for responses to be scored. Items were phrased carefully, simply, and unambiguously, and were numbered, as this is a familiar questionnaire style and recommended in psychometric texts (Rust & Golombok, 1989). The use of double-negatives was avoided, and each item dealt with just one issue. Social desirability was discouraged by careful phrasing, by instructions to avoid spending too long on any one question, and by emphasizing that the first response is usually the best. The questionnaire was presented in two parts, A (7 items) and B (56 items), respectively relating to respondents' experiences before and during the mediated environment. Comprehensive instructions explained the terms used, the layout, and how to select a response, and provided an assurance of confidentiality to increase compliance and honesty. A number of potentially important background variables were also collected. These included basic demographics (such as age, sex, occupation), education, self-rated experience with computers and frequency of computer games playing, and self-rated knowledge of VR, TV/film production, and stereoscopic image presentation using polarized spectacles. This section requires some modification for use in non-UK labs, because educational status is tailored to UK standards. Alternatively, the background information can be replaced with background variables that are of interest to a particular lab.

2.2 Samples and Procedure

The ITC-SOPI was administered to a total of 604 people following an experience with a mediated environment. The sample was divided into six subsamples, each of which experienced a different level of physical immersion (for example, 3-D versus 2-D, ability to control aspects of the environment versus no control, large screen versus small screen, surround sound versus stereo) and different display contents.

Respondents were recruited from an IMAX cinema, presenting 2-D and 3-D (viewed using polarized glasses) films on a 15m × 21m screen, $n = 22$ and 196, respectively; a well-known cinema chain in the United Kingdom, presenting 2-D films on general release, $n = 238$; Goldsmiths College Students' Union (GCSU) film night, where VHS films are projected onto a 2m × 1.5m (approx.) display screen with fairly degraded audio/visual fidelity, $n = 26$; and two experimental settings based in the Psychology Department at Goldsmiths College. One of these consisted of participants viewing a video short (<30 min.) on a 28 in. color TV ($n = 73$) before completing the ITC-SOPI. The other experimental setting entailed participants playing a computer-generated racing game on a consumer videogame console ($n = 49$). The console was connected to the Platform for Immersive Television (P*i*T, a controlled test environment with a 28 in. color TV and set to mono audio output: see figure 1). Participants navigated the car around the track using an unsophisticated gamepad. An incentive was offered for questionnaire completion (free prize draw entry, free "two tickets for the price of one" vouchers, cash payment of £3, or course credit (psychology undergraduates only)).

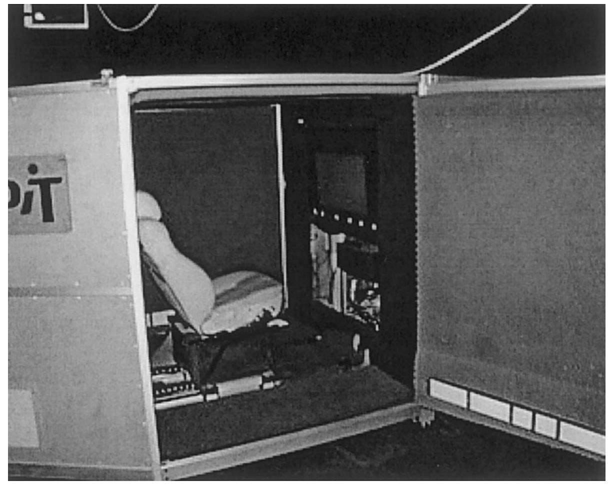

Figure 1. *Venue example: The Platform for Immersive Television (PiT) for computer games console respondents.*

The content was varied to avoid confounding the results with a specific type of content. In this way, the stability of the relationships between questions, irrespective of content, could be examined. All respondents rated just one presentation. There were two presentations for IMAX 3-D, both fictitious; one for IMAX 2-D (a documentary film); eleven different films for the cinema condition (including one animated film); two films for GCSU; six video shorts, comprising two animated stimuli, football highlights, a black-and-white French film (no subtitles and eclectic editing), and two crime reconstructions from a popular British television program; and one stimulus for the computer games sample (F1 Racing). Although the subsamples had a broad range of media content in terms of different themes and stories, in terms of media form, noninteractive, photorealistic displays were heavily weighted in the sample. These are potential shortcomings that we address in the discussion.

Fifty-one percent of the sample were male, and respondents were aged between 9 and 73 years (mean age of 29 years; s.d. = 11.24).

3 Results

3.1 Principal Axis Factoring

Data for the 63 items were entered into a principal axis factoring (PAF) analysis. (For an introduction to factor analysis, see Tabachnick and Fidell (1989, chapter 12)). In brief, factor analysis is a data-reduction technique that summarizes patterns of correlations by revealing groups of correlated items, which are called *factors*. Each factor accounts for a proportion of the variance across all items in the data set. This proportion is represented by a numerical term, an eigenvalue. The first factor will always account for the largest proportion of variance in a data set (that is, it will have the largest eigenvalue). Eigenvalues decrease with each successive factor. Factor loadings represent correlations between each item and a factor. The magnitude of the correlations therefore indicate the weightings of variables on each factor. In a good factor analysis, a variable will load highly on just one factor, although variables will sometimes cross-load. Ideally, a minimum of three variables will mark a factor.

Pairwise deletion for cases with missing values was used to make full use of the data set. This was repeated with the more conservative listwise deletion. Both produced broadly similar structures and factor loadings. Preliminary checks (such as the Kaiser-Meyer-Olkin measure of sampling adequacy) indicated that the data were suitable for PAF. Furthermore, examination of the intercorrelation matrix for all variables revealed a large number of significant correlations ($p \leq 0.001$) indicating good factorability of the data set.

A number of criteria were employed to determine the number of factors to extract. First, factors with initial eigenvalues greater than 1 were considered. According to Tabachnick and Fidell (1989), this figure is "usually somewhere between the number of variables divided by 3 and the number of variables divided by 5" (p. 635). For 63 items, somewhere between 13 and 21 would therefore be anticipated. In fact, twelve factors satisfied this criterion (16.47, 4.81, 2.61, 2.55, 2.00, 1.84, 1.48, 1.36, 1.27, 1.18, 1.06, 1.05). With 40 or fewer variables and a large sample size, this criterion is usually a good index, but, in other circumstances, the criterion of eigenvalues greater than 1 can over- or underestimate the number of factors (Tabachnick & Fidell, 1989). In this instance, with 63 items, the number of factors was likely to be overestimated.

A second criterion, of examining "elbows" in the scree plot (a plot of the eigenvalues), suggested either a two-, three-, four-, or five-factor solution. A third criterion was based on comparing the number and spread of "marker" variables for each factor solution. (See table 2.) We use two definitions of marker variables. First, they are defined as any variable with a factor loading above 0.3 on any factor in the rotated factor matrix (Watson et al., 1995). A more stringent definition is that a variable's primary loading must be 0.2 greater than any cross-loading (Bedford, 1997), in addition to having a primary loading above 0.3. When comparing different factor solutions, the greater the number of marker variables per factor coupled with a relatively even distribution of markers that meet the more stringent criterion indicates a better solution for the data. Note, however, that fewer marker variables are anticipated with each successive factor that is extracted in any solu-

Table 2. *Marker Variables on Each Factor for the Five Different Factor Structures*

Number of factors in solution	1	2	3	4	5
1	50				
2	47 (30)	26 (10)			
3	36 (17)	29 (8)	17 (4)		
4	28 (14)	24 (13)	16 (6)	8 (6)	
5	27 (12)	24 (12)	13 (4)	7 (6)	6 (0)

Values are based on each of two criteria: (i) Watson et al.'s (1995) criterion of variables that load higher than 0.3 in the rotated factor matrix, and values in parentheses are based on (ii) Bedford's (1997) criterion of variables that load higher than 0.3 and where the primary loading is more than 0.2 greater than any cross-loading.
No value is available for the one-factor solution using criterion (ii) because there was no other factor with which to compare loadings.

Table 3. *Items that Failed to Load Significantly (≥0.3) on any of the Four Factors which were Deleted from the Revised ITC-SOPI*

	Loading			
ITC-SOPI item	F1	F2	F3	F4
I was aware of the real world	−0.23	−0.15	−0.03	−0.11
I wanted to see more of the space in the displayed environment than I was able to	0.18	0.23	0.07	0.19
I found it easy to forget that I was watching a display	0.19	0.23	0.25	0.19
I had the best viewpoints	0.20	0.19	0.19	−0.05
The temperature of the real world distracted me	0.05	0.03	−0.12	0.26
I was distracted by the quality of the technology	0.15	0.07	−0.17	0.29
I wanted to make specific sounds louder or softer	0.20	0.11	−0.05	0.25
I felt I knew what was going to happen next	0.02	−0.02	−0.11	0.03

tion. The five-factor solution was rejected because, using Bedford's more stringent criterion, there were no marker variables on the fifth factor, as shown in table 2. Although the two-, three-, and four-factor solutions each fared reasonably well on the more stringent criterion, the four-factor solution made the most conceptual sense. Cumulatively, the four factors accounted for 38.3% of the variance in the data.

The four factors were subjected to varimax rotation, which aims to achieve "simple structure" (Thurstone, 1947). Simple factor structures are easier to interpret as they avoid ambiguities associated with correlated factors. A similar solution to that we report here was obtained using oblique rotation, but, for our preliminary data reduction purposes, we report the varimax solution.

3.1.1 Factor Structure. The first factor accounted for 14.2% of the variance and was defined by high loadings from items such as "I had a sense of being in the scenes displayed," "I felt I was visiting the places in the displayed environment," and "I felt that the characters and/or objects could almost touch me." These items indicate a sense of physical placement in the mediated environment, and interaction with and control over parts of the mediated environment. This factor seems to encapsulate the definitions of presence discussed in section 1—a sense of being there—and was consequently labeled "Sense of Physical Space."

The second factor explained 11.1% of the variance and had high loadings from items including "I felt involved (in the displayed environment)," "I enjoyed myself," and "My experience was intense." These items suggest a tendency to feel psychologically involved and to enjoy the content. This factor was consequently labeled "Engagement."

The third factor explained 7.6% of the variance and was characterized by questions including "The content seemed believable to me," "The displayed environment seemed natural," and "I had a strong sense that the characters and objects were solid." These variables indicate a tendency to perceive the mediated environment as lifelike and real, and the factor was subsequently labeled "Ecological Validity."

The fourth and final factor explained 5.4% of the variance and was almost exclusively characterized by the variables describing adverse physiological reactions such as "I felt dizzy," "I felt nauseous," "I felt I had a headache," and "I had eyestrain." This factor was labeled "Negative Effects."

Although all items/questions were generated or selected from previous presence measures for their putative ability to tap elements of presence, eight questions failed to load significantly (≥0.30) on any of the factors. (See table 3.) These were discarded for the revised version of the ITC-SOPI.

3.2 Reliability

First, to examine the stability of the items on each of the four factors, the data set was randomly divided into two subsamples (group 1: $n = 325$; group 2: $n = 279$), while ensuring that each of the viewing conditions (IMAX 2-D, 3-D, Cinema GCSU film night, video shorts, and computer game) was adequately represented within each subsample. Most of the items showed the same loadings as those obtained in the large sample. In an attempt to further reduce the number of items in the ITC-SOPI, six items that were conceptually or statistically inconsistent were excluded from the revised questionnaire.

Second, internal reliability coefficients (Cronbach's alpha) were computed for each of the four factors. Items that failed to load above 0.3, or that were deemed conceptually or statistically inconsistent following the structure reliability check described above, were excluded. Alphas were very good: Sense of Physical Space (21 items) = 0.94; Engagement (13 items) = 0.89; Ecological Validity (6 items) = 0.76; Negative Effects (8 items) = 0.77. Items that reduced the alpha coefficient (two for Sense of Physical Space, one for Ecological Validity, and two for Negative Effects) were excluded from the revised ITC-SOPI.

3.3 Validity: Media Format

Preliminary validation of the ITC-SOPI was performed by comparing mean scale scores computed for each factor across the different media formats (samples). The results were very positive particularly for the first factor given that several important issues, such as content, were not controlled. (See figure 2a–d.) Clearly, a more stringent validity test would compare matched content across the different media formats.

Sense of Physical Space showed sensitivity to media form: IMAX 3-D had the highest score, followed by interactive computer game, IMAX 2-D, cinema, GCSU, and video shorts. (See figure 2a.) A one-way analysis of variance (ANOVA) revealed a significant difference in Sense of Physical Space scores across the six media samples ($F(5,603) = 75.77$; $p < 0.001$). Post-hoc tests (Bonferroni) indicated that ratings for Sense of Physical Space for all the media formats were significantly different from one another ($p < 0.05$) with the exceptions of IMAX 2-D versus computer games, video shorts versus GCSU, and cinema versus GCSU.

Engagement also demonstrated some sensitivity to media sample: mean scores were highest for IMAX 2-D, followed by IMAX 3-D, computer game, GCSU, cinema, and video shorts. A one-way ANOVA revealed a significant difference between the six media samples ($F(5,603) = 20.81$; $p < 0.001$). Post-hoc tests (Bonferroni) indicated that ratings for Engagement were significantly different ($p < 0.05$) between video shorts and every other condition, between IMAX 2-D versus cinema, and IMAX 3-D versus cinema. (See figure 2b.)

For Ecological Validity, mean scores were highest for IMAX 2-D, followed by IMAX 3-D, video shorts, cinema, GCSU, and computer game. A one-way ANOVA showed a significant difference between the six media samples ($F(5,603) = 23.61$; $p < 0.001$). Post-hoc tests (Bonferroni) indicated that ratings for Ecological Validity for all the media formats were significantly different from one another ($p < 0.05$) with the exceptions of IMAX 3-D versus video shorts, cinema versus computer game, cinema versus GCSU, and computer game versus GCSU. (See figure 2c.)

Finally, for Negative Effects, IMAX 2-D had the highest mean score, followed by IMAX 3-D, computer game, video shorts, cinema, and GCSU. A one-way ANOVA revealed a significant difference between the six conditions ($F(5,603) = 12.03$; $p < 0.001$). Specifically, post-hoc tests (Bonferroni) revealed significant differences ($p < 0.05$) between IMAX 2-D versus cinema, IMAX 2-D versus GCSU, IMAX 3-D versus video shorts, IMAX 3-D versus cinema, IMAX 3-D versus GCSU; all other pairs were not significant. (See figure 2d.)

3.4 Revised ITC-SOPI

A total of 44 items were retained following the analyses detailed above. Items were dropped that either: (i) failed to load above 0.3 on any factor (eight items);

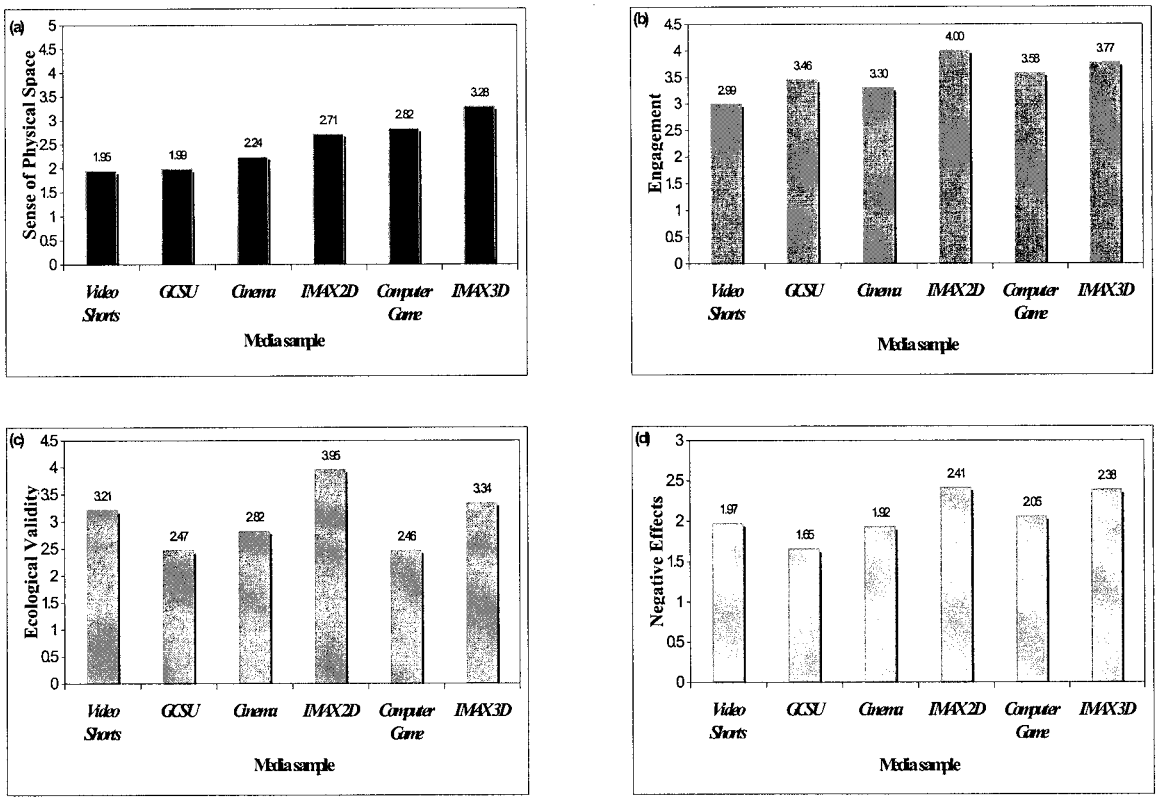

Figure 2. *Mean scores for each of the media samples: (a) Sense of Physical Space; (b) Engagement; (c) Ecological Validity; and (d) Negative Effects.*

(ii) showed unstable factor loadings when the sample was randomly split and/or were deemed of low importance or relevance to their factor (six items), or (iii) elevated the internal consistency of their respective factors when deleted in the alpha check (five items).

3.5 Factor Intercorrelations

Scores were produced for each of the four factors by calculating the mean of their respective items (based on the revised ITC-SOPI described above, where Sense of Physical Space = 19 items; Engagement = 13 items; Ecological Validity = 5 items; and Negative Effects = 6 items). Pearson's correlations were run between all the resultant scale totals ($n = 604$). Each of the first three "positive presence-related" scales were strongly positively intercorrelated (Sense of Physical Space and Engagement: $r = 0.62$; $p < 0.001$; Sense of Physical Space and Ecological Validity: $r = 0.52$; $p < 0.001$; Engagement and Ecological Validity: $r = 0.45$; $p < 0.001$). Negative Effects did not correlate with Engagement ($r = -0.05$; n.s.) or with Ecological Validity ($r = 0.06$; n.s.), but did correlate with Sense of Physical Space ($r = 0.24$; $p < 0.001$).

4 Discussion

The factor structure identified here, based on responses to the ITC-SOPI across a range of predominantly noninteractive media formats, is similar to that produced by other researchers who have attempted to elucidate the structure of presence using more-limited samples (for example, responses from VE users or 3-D computer games players). Informal reports in both VE (Usoh et al., 1999) and TV-based studies (Freeman & Avons, 2000) suggest that there are distinct similarities in respondents' experiences of "being there" elicited by different display media.

Each of the four factors is likely to be determined by the interaction between media form, media content, and user characteristic variables. In this regard, it is not possible to draw firm conclusions based on the present data set, which intended to include as much variety in these three broad determinants as possible to enhance the generalizability of the measure. Further studies are required that measure the effects of manipulations of specific determinants of presence on the ITC-SOPI factors. However, we speculate briefly on possible relations between determinants of presence and each of the factors we identified in our current dataset, based on results from the current data set and the constituent items of each factor.

The first factor in our data set, Sense of Physical Space, contains two of the three items used in simple post-test presence scales that have previously been considered to tap key elements of presence (such as the scales of Slater et al., 1994). It is of note that, although these items had primary loadings on Sense of Physical Space, they also cross-loaded on Engagement and Ecological Validity. Presence is likely to be related not only to a user's sense of being located within a spatially contiguous physical environment but also to his/her personal evaluation of the appeal and the naturalness/believability of the content within the displayed environment.

We propose that the prime determinants of the Sense of Physical Space factor are, in the terminology we use in the introduction to this paper, media form variables. Some support for this assertion is given by the preliminary validity analysis that found that users of high-fidelity presentations (such as IMAX 3-D) gave higher Sense of Physical Space ratings than did users of lower-fidelity presentations (such as GCSU). Interestingly, although the computer-games console condition consisted of a relatively small field of view and was low in photorealism, respondents produced relatively high Sense of Physical Space ratings. This suggests that the ability to physically control and manipulate aspects of the displayed environment (even using unsophisticated control devices) enhances the sense of being physically located in that environment. Media form variables may therefore interact in compensatory ways, similar to the isopresence equivalence classes suggested by Ellis (1996). Furthermore, it is of note that these results were obtained when content was varied both within and across the samples piloted.

The second factor in our data set, Engagement, provides a measure of a user's involvement and interest in the content of the displayed environment, and their general enjoyment of the media experience. That this factor emerged indicates that attention and involvement are important elements of a user's evaluation of a media experience (Barfield & Weghorst, 1993; Draper et al., 1998; Witmer & Singer, 1998). The content of the presentation inevitably influences ratings of Engagement. Indeed, one of the ITC-SOPI items directly probes respondents about how appealing they found the content. Other questions that comprise this scale relate to arousal and emotionality. These are likely to be influenced by the media content, but also intensified by the media form. For instance, a fast-paced rally car scene (with first-person perspective) is likely to be arousing in itself, but it is likely to be more arousing when presented on a large screen with surround sound and motion feedback. By this rationale, Engagement may be determined both by media content and media form variables.

Questions that comprise our third factor, Ecological Validity, relate to the believability and realism of the content and the naturalness and solidity of the environment. The number, extent, and consistency of sensory stimulation (media form variables) are therefore likely to enhance perceived naturalness and, in turn, increase ratings on this scale. Indeed, this is somewhat supported by the pattern of scores across the media samples piloted. Again, IMAX presentations received the highest ratings, in spite of the entirely fictitious contents of IMAX 3-D. This suggests that the greater the immersive properties of the media form, the less influence that content has on the presentation's believability, realness, and naturalness. Differences in the media form variable, photorealism, should also produce differences on this scale. Indeed, the computer game sample received the lowest ratings compared to the other media samples on Ecological Validity.

Our fourth factor, Negative Effects, is less related to our first three factors than they are to each other. Witmer and Singer (1998) report that presence and simulator sickness are inversely related. Whilst in the current study Negative Effects was not strongly correlated (positively or negatively) with Engagement or Ecological Validity, it was significantly but modestly (and positively) related to Sense of Physical Space. Headache, eyestrain, tiredness, and other negative effects may be associated with the media form. For instance, some respondents in the IMAX 3-D sample reported (in additional comments) that the polarized glasses were uncomfortable. Indeed, IMAX 3-D respondents gave the second-highest ratings on this scale. IMAX 2-D respondents, on average, gave the highest Negative Effects ratings. This sample had no variation in content; all respondents viewed *Everest,* a documentary following the ascent of a team of climbers of the world's highest mountain, and this may have produced higher than average negative effects such as dizziness and disorientation. Content is also likely to affect self-reported negative effects more generally by interacting with user characteristics (such as personal content preferences). For instance, if a given content is perceived as boring, corresponding ratings of tiredness, or even headaches, may result.

Schubert et al.'s (1999) three presence-related components ("Spatial presence", "Involvement", and "Realness [comparability to the reality]") have striking parallels with the first three factors identified in the present study (Sense of Physical Space, Engagement, and Ecological Validity). Witmer and Singer's (1998) first cluster, "Involved/Control," is comparable to an amalgamation of our Sense of Physical Space and Engagement factors. Indeed, Sense of Physical Space and Engagement were found to correlate strongly and significantly, lending support to Witmer and Singer's (1998) notion that "Immersion" (Sense of Physical Space) and "Involvement" (Engagement) are interdependent. Their second and third clusters, "Natural" and "Interface quality" are akin to Ecological Validity as identified here. Our Negative Effects factor is not represented in either Witmer and Singer's or Schubert et al.'s structures, because they did not include questions relating to adverse physiological effects of exposure to media.

It is less clear how the factors we have obtained relate to Kim and Biocca's (1997) two-factor solution. It is likely that the limited number of items (eight) that Kim and Biocca used did not provide a comprehensive coverage of all the possible facets of presence that were incorporated into the ITC-SOPI.

From a theoretical perspective, our data are consistent with Slater and Wilbur's (1997) notions of the influence of immersion on presence. They argue that high presence is associated with the extent to which an individual is provided with an inclusive, extensive, surrounding, and vivid display. This is not to suggest that our questionnaire measures objective physical properties of the technologies used; it explicitly measures subjective experience. However, our preliminary validation check demonstrates that our first factor, Sense of Physical Space, is sensitive to variation in objective system parameters. It is our view that these parameters should be treated as independent variables where researchers have an interest in their effects on presence. In addition, Sense of Physical Space and Engagement correspond closely with two factors that Witmer and Singer (1998) identify as essential to experience presence (which they term "Immersion" and "Involvement").

It is not clear whether the four factors we have identified all contribute to a sense of presence, and, if they do, whether this occurs additively or in a more complex manner. We currently recommend that each ITC-SOPI scale be analyzed separately, as we expect them to be differentially sensitive to manipulations of particular determinants of presence.

The preliminary results reported in this paper are encouraging but require replication with new samples and evenly distributed types of content. Clearly, one limitation of the ITC-SOPI is that it has not been piloted on a broad range of what we term "interactive" displays (that is, those that include a control/manipulation device). However, the revised ITC-SOPI has now been distributed to a number of VE labs across the world. Confirmatory factor analysis of these new data will enable a test of the stability of both the factor structure and its constituent items. A full exploration of the interrelationships between the factors will then be undertaken. In addition, guidelines for scoring ITC-SOPI responses will be developed. A more rigorous validation of the ITC-SOPI is also in progress, using experimental designs controlling for factors such as novelty and content. Finally, corroborative evidence will be sought through behavioral and physiological measures of presence collected concurrently with the revised ITC-SOPI.

As we discussed in the introduction to this paper, a standard cross-media presence measure will constitute a useful tool for the presence research community. It will enable the comparison of results both within and across independent research laboratories. Across labs, variations in media content will clearly affect presence ratings; there is, after all, likely to be substantial variation in users' attitudes toward different contents, in turn, influencing presence. This is an issue that is likely to affect any presence measure. Nonetheless, it is reassuring that the first ITC-SOPI factor, Sense of Physical Space, demonstrated sensitivity to variations in media form in spite of variation in content. A generic tool will enable researchers to investigate in more detail the relative importance of the determinants of presence, including aspects of media form, media content, and user characteristics.

Current questionnaire measures have not been developed for the evaluation of presence across a range of media because their piloting procedure has involved very limited samples. Although the present limitations of the ITC-SOPI in this regard are acknowledged, it has been designed for general utility without emphasizing system features of any particular medium. However, it is recommended that the ITC-SOPI be used in conjunction with other questionnaires tailored to lab-specific requirements. These might range from questions focusing on particular features of media content, such as perceived relationships with characters in a mediated environment, to the usability of specific system devices. Subject to further validation, we expect the ITC-SOPI to offer great utility to presence researchers using very different media systems ranging from standard broadcast displays to advanced, fully immersive VEs.

Acknowledgments

The generous support of Dr. Nick Lodge and Dr. David Harrison (Technology Group) of the UK Independent Television Commission to the Immersive TV Project at Goldsmiths College, University of London is gratefully acknowledged. Discussions with Cath Dillon (Goldsmiths College, London) and Dr. Matthew Lombard (Temple University, USA) were also useful in the early stages of the questionnaire development. The authors would also like to acknowledge with thanks their partners from the European Commission's ACTS TAPESTRIES project (1996–1999), in particular Dr. S.E. Avons (University of Essex) and Dr. Wijnand IJsselsteijn (IPO Center for User-System Interaction, NL), with whom several of the ideas behind the research reported here were originally discussed. Finally, we are very grateful to the media companies whose customers completed the ITC-SOPI.

References

Barfield, W., Baird, K. M., & Bjorneseth, O. J. (1998). Presence in virtual environments as a function of type of input device and display update rate. *Displays, 19,* 91–98.

Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. In G. Salvendy & M. Smith (Eds.) *Software and hardware interfaces* (pp. 699–704). Amsterdam: Elsevier.

Barfield, W., Zeltzer, D., Sheridan, T. B., & Slater, M. (1995). Presence and performance within virtual environments. In W. Barfield & Furness, T. A. (Eds.). *Virtual envi-*

ronments and advanced interface design (pp. 473–541) Oxford: Oxford University Press.

Bedford, A. (1997). On Clark-Watson's tripartite model of anxiety and depression. *Psychological Reports, 80*(1), 125–126.

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors, 40,* 354–375.

Ellis, S. R. (1996). Presence of Mind: A reaction to Thomas Sheridan's "Further Musings on the Psychophysics of Presence." *Presence: Teleoperators and Virtual Environments, 5,* 247–259.

Fontaine, G. (1992). The experience of a sense of presence in intercultural and international encounters. *Presence: Teleoperators and Virtual Environments, 1,* 482–490.

Freeman, J. (2000). *Subjective and Objective Approaches to the Assessment of Presence.* Unpublished doctoral dissertation, University of Essex.

Freeman, J., & Avons, S. E. (2000). Focus group exploration of presence through advanced broadcast services. *Proceedings of the SPIE, Human Vision and Electronic Imaging V,* 3959–3976.

Freeman, J., Avons, S. E., Pearson, D., & IJsselsteijn, W. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments, 8,* 1–13.

Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., & IJsselsteijn, W. (2000). Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. *Presence: Teleoperators and Virtual Environments, 9*(2), 149–164.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and Virtual Environments, 1,* 262–271.

Hendrix, C., & Barfield, W. (1996a). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments, 5,* 274–289.

———. (1996b). The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments, 5,* 290–301.

IJsselsteijn, W., de Ridder, H., Hamberg, R., Bouwhuis, D., & Freeman, J. (1998). Perceived depth and the feeling of presence in 3DTV. *Displays, 18,* 207–214.

IJsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. (2000). *Presence: Concept, determinants and measurement.* Proceedings of the SPIE, Human Vision and Electronic Imaging V, 3959, 520–529.

Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer Mediated Communication* [On-line serial], 3(2). Available: http://www.ascusc.org/jcmc/vol3/issue2/kim.html.

Kline, P. (1993). *Personality: the psychometric view,* London: Routledge.

Lodge, N. K. (1999). *Being part of the fun—Immersive television! Keynote address at the Conference of the Broadcast Engineering Society (India),* New Delhi, February 1999.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication,* [On-line serial] 3(2), Available: http//www.ascusc.org/jcmc/vol3/issue2/lombard.html.

North, M. M., North, S. M., & Coble, J. R. (1997). Virtual environments psychotherapy: A case study of fear of flying disorder. *Presence: Teleoperators and Virtual Environments, 6*(1), 127–132.

Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement.* London: Pinter Publishers.

Rust, J., & Golombok, S. (1989). *Modern psychometrics: The science of psychological assessment.* London: Routledge.

Schubert, T. W., Friedmann, F., & Regenbrecht, H. T. (1999). *Decomposing the sense of presence: Factor analytic insights.* Paper presented at the Second International Workshop on Presence, University of Essex, Colchester, April 1999 [on-line]. Available: http://www.uni~jena.de/~sth/vr/insights.html.

Sheridan, T. B. (1992a). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments, 1,* 120–125.

———. (1992b). Defining our terms. *Presence: Teleoperators and Virtual Environments, 1,* 272–274.

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments, 8,* 560–565.

Slater, M., Steed, A., McCarthy, J., & Maringelli, F. (1998). The influence of body movement on presence in virtual environments. *Human Factors, 40,* 469–477.

Slater, M., & Usoh, M. (1994). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments, 2,* 221–233.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 3,* 130–144.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 6,* 603–616.

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics.* (2nd ed.). New York: Harper Collins Publishers.

Thurstone, L. L. (1947). *Multiple Factor Analysis.* Chicago: University of Chicago Press.

Usoh, M., Arthur, K., Whitton, M., Bastos, R., Steed, A., Brooks, F., & Slater, M. (1999). *The visual cliff revisited: A virtual presence study on locomotion*. Paper presented at the Second International Workshop on Presence. University of Essex, Colchester, April 1999 [on-line]. Available: http://www.cs.ucl.ac.uk/staff/m.usoh/bt.html.

VIRTUE (2000). European Commission, Information Society Technologies Project [on-line]. Available: http://www.cordis.lu/ist/projects/99-10044.htm.

Watson, D., Clark, L. A., Weber, K., Assenheimer, J. S., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal Psychology, 104*(1), 15–25.

Welch, R. B., Blackmon, T. T., Liu, A., Mellers, B. A., & Stark, L. W. (1996). The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence: Teleoperators and Virtual Environments, 5,* 262–273.

Welch, R. B. (1999). How can we determine if the sense of presence affects task performance? *Presence: Teleoperators and Virtual Environments, 8,* 574–577.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7,* 225–240.