Physiological Measures of Presence in Virtual Environments

Michael Meehan, Brent Insko, Mary Whitton, Dr. Frederick P. Brooks, Jr.
University of North Carolina
Chapel Hill, North Carolina, U.S.A.
(919) 962-1700
(meehan, insko, whitton, brooks)@cs.unc.edu

1 Presence and virtual environments

Virtual environments (VEs) are the most advanced human-computer interfaces yet developed. Researchers, by the development of new methods, theories, and technologies, have endeavored to make effective VEs. The definition of effectiveness changes based on the application of the VE. For flight simulators, training transfer is important. For architectural walkthroughs, accurate perception of space is important. For treatment of phobias and post-traumatic stress disorders, presence – evoking in patients the feeling that they are near the source of their phobia or stress – is important [Hodges, 1994]. It is on this last concept, presence, that this paper focuses.

Rothbaum and Hodges' VE system for graded exposure treatment of acrophobia strives to bring patients near the source of their phobias [Hodges, 1994]. They state that "the user's sense of presence is the defining factor in the [successful treatment of acrophobia]." We believe this is true for all phobia treatment systems: the system must evoke presence in order to work. Such systems are useful as they allow much of the effectiveness of *in vivo* exposure with the safety, convenience, and reduced cost of in-office therapy [Hodges, 1995]. To ensure the systems evoke presence in users, developers endeavor to build the best VEs possible: stereo portrayal (as opposed to mono) in the headmounted display, realistic models and lighting, low lag, high frame rate, etc.

VE developers, though, have limited time and resources, and these limitations force system design choices. When making these choices, designers must ensure that the users still have a compelling experience – that the VE still evokes a sense of presence. So they need to know what is important for evoking presence in

VEs: Is increasing the frame rate more important than stereo display in the headmounted display (HMD)? Is it more important to have lower lag or a richer model? Is pixel density (resolution) more important than field-of-view? Are dynamic shadows the key to making a VE convincing?

The answers to these questions may be different for every person and even for the same person at different times. We believe, however, that a broad suite of thoughtfully constructed studies investigating the effects of varying VE system parameters (lag, frame rate, realism) on presence would reveal rules of thumb for *what evokes presence* for the general population. To find these rules, however, we need to be able to measure presence, and the measure must be reliable, valid, multi-level sensitive, and objective. We investigated physiological reactions as such measures.

2 Measuring presence

The concept of *presence* is difficult to define, and becomes even more so when one tries. Most of the definitions presented have discussed the concept of "being there" in a virtual environment. We define presence in this spirit as *perceiving stimuli as one would perceive stimuli from the corresponding real environment*. The stimuli that the user perceives come from the VE in our experiments.

Since presence is a subjective condition, it has most commonly been measured by self-reporting, either during the VE experience or afterwards by questionnaires. There has been vigorous debate as to how to define presence and how to best measure it [Barfield, 1995; Ellis, 1996; Freeman, 1998; IJsselsteijn, 1998; Lombard, 1997; Lombard, 1999; Meehan, 2000a; Meehan, 2000b; Regenbrecht, 1997; Schubert, 1999; Sheridan, 1996; Slater, 1999; Witmer, 1998]. In their investigations, researchers have yearned for a measure that is

Reliable – produces repeatable results, both from trial to trial on the same subject and across subjects;

Valid – measures subjective presence, or at least correlates well with established subjective presence measures;

Multi-level sensitive – is capable of distinguishing multiple levels of presence; and

Objective – is well shielded from both subject bias and experimenter bias.

We attempted to create such a measure and report our findings here. We investigated physiological reactions as measures of presence over multiple exposures, both on a single day and over multiple days. We also investigated the measures in multiple presence conditions using *passive haptics* (a rough physical model corresponding to the VE) and multiple *frame rates* (the number of times per second that the image in the headmounted display is updated to reflect the user's current position). The results of all three studies are discussed below. First, we discus the findings in terms of investigating physiological reaction as a measure of presence. Later, we regroup the findings and summarize the findings in the three studies: Effect of Multiple Exposures on Presence, Effect of Passive Haptics on Presence, and Effect of Frame Rate on Presence.

Our thesis is that

To the degree that a virtual environment evokes presence (as defined above), it will evoke physiological responses similar to those evoked by the corresponding real environment, and greater presence will evoke a greater response. Hence, these physiological responses can serve as reliable, valid, multi-level sensitive, and objective measures of presence.

We used a VE that simulates a danger-of-falling, stress-inducing environment (see Figure 1) and selected certain physiological responses that were easy to measure from the hands and chests of the subject and have documented responses to this stress: heart rate, skin conductance, and skin temperature. Heart rate and skin conductance are known to increase and skin temperature decrease with exposure to heights and other stressors. For example, Emmelkamp and Felten reported on nineteen acrophobic patients' heart rate reactions to climbing "as high as they could" on a fire escape (with a hand rail), waiting one minute, and looking down. Subjects ascended to the second landing on average. The average heart rate increase for subjects was 13.4 beats / minute [Emmelkamp, 1985]. Our subjects were non-phobic, so we would expected their heart rate reactions to be lower, but in the same direction. See [Andreassi, 1995; Guyton, 1986] for more discussion on physiological reaction.

2.1 The Measures

Physiological measures of presence. We constructed three physiological measures based on differences between stress reactions and normal values. These were defined so that they should all increase with increased presence. That is, if there was more presence, there should have been more physiological reaction to the Pit Room (Figure 1), and these measures should each increase:

ΔHeart Rate = Mean Heart Rate $_{\text{Pit Room}}$ - Mean Heart Rate $_{\text{Training Room}}$

ΔSkin Conductance = Mean Skin Conductance $_{\text{Pit Room}}$ - Mean Skin Conductance $_{\text{Training Room}}$

ΔSkin Temperature = Mean Skin Temperature $_{\text{Training Room}}$ -Mean Skin Temperature $_{\text{Pit Room}}$

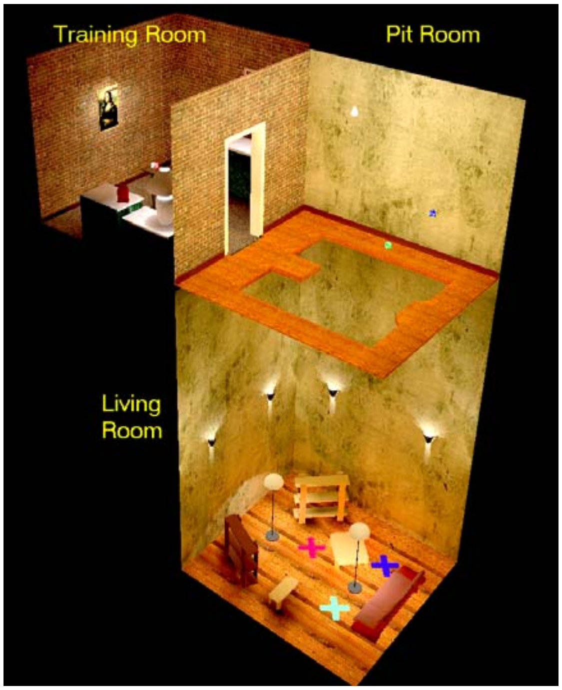


Figure 1. Side view of the virtual environment. Subjects start in the Training Room and later enter the Pit Room.

Reported Presence and Reported Behavioral Presence. To measure reported presence, we used a modified version of the University College London (UCL) Presence Questionnaire [Usoh, 1999]. The UCL questionnaire contains seven questions that measure *presence* (Reported Presence), three questions that measure *behavioral presence* (Reported Behavioral Presence) – Did the user report acting as he would in a similar real environment? – and three that measure *ease of locomotion* (Reported Ease of Locomotion) – Did the user report that it was easy and natural to move about in the virtual environment? Reported Ease of Locomotion is not a measure of presence and therefore we do not discuss it extensively here. Responses for each question were on a scale of 1 to 7.

Observed Behavioral Presence. We videotaped all the subject sessions and then, from the tapes, scored presence depending upon various characteristic behaviors including *taking baby steps, testing the edge with the foot,* etc. During our sessions, technical problems caused some sessions tapes to be unusable due to lighting or due to difficulties with the recording equipment. In the Passive Haptics study, these technical problems left us with data for only 31 of 52 subjects. We defined the movements to be scored beforehand and had only one experimenter scoring for each study. A more detailed investigation of this measure should ensure a complete data set by ensuring that all sessions are properly recorded. A detailed investigation should also use multiple scorers and investigate the reliability among these scorers.

Figure 2. View of the 20' pit from the edge of the diving board.

2.2 The Experiment

The environment. The environment shown in Figures 1 and 2 consisted of 10,000 polygons and 41 megabytes of texture maps. This environment was used in the Frame Rate study. A similar environment was used for the Passive Haptics and Multiple Exposures studies. It had 20,000 polygons and 50 megabytes of texture. All environments were rendered in stereo on one Infinite Reality 2 pipe of an SGI Reality Monster. The head-mounted display was a Virtual Reality 8 with 640x480 tri-color pixel resolution in each eye. Users walked about in an 18'x 32' space, tracked with a high-accuracy, very-low-lag University of North Carolina Hi-Ball optical tracker [Ward, 1992; Welch, 1997]

Experimental procedures. In the Multiple Exposures study, 10 subjects (average age 24.4; σ = 8.2; 7 female, 3 male) were trained to pick up books and move about in the Training Room – at which time a physiological baseline was taken. Subjects then carried a book from the Training Room and placed it on a chair on the far side of the Pit Room. After that, they were instructed to return to the Training Room. The subjects performed this task three times per day on four separate days. In the Multiple Exposures study, we investigated the hypothesis that the presence-evoking power of a VE declines with multiple exposures. ΔHeart Rate was not successfully measured in this study due to problems with the sensor. We excluded subjects who had experienced VEs more than three times from all studies.

In the Passive Haptics study, 52 subjects (average age 21.4; σ = 4.3; 16 female, 36 male) reported on two days. On one day, a subject experienced the VE with the 1.5-inch wooden ledge - on the other day, without the ledge. Subjects were counterbalanced as to the order of presentation of the ledge. Subjects performed all exposures to the VE wearing only thin sock-like slippers. The task was the same as in the Multiple Exposures study except subjects were instructed to walk to the edge of the wooden platform, place their toes over the edge, and count to ten before dropping the book on the chair on the far side of the Pit Room. In the Passive Haptics study, we investigated the hypothesis that the 1.5-inch wooden ledge increased the presence-evoking power of the VE. Figures 3 and 4 show a subject standing over the visual ledge with and without the 1.5-inch wooden ledge present.

Figure 3. Subject standing on 1.5-inch wooden ledge.

Figure 4. Subject standing over the visual ledge with no 1.5-inch wooden ledge present.

The Frame Rate study had 33 participants (average age 22.3; $\sigma = 3.6$; 8 female, 25 male). Subjects entered the VE four times on one day and were presented the same VE with a different frame rate each time. The four frame rates were 10, 15, 20, and 30 frames-per-second (FPS). Subjects were counterbalanced as to the order of presentation of the four frame rates. Subjects were trained to pick up and drop blocks in the Training Room and then carried a red block to the Pit Room and dropped it on a red X-target on the floor of the Living Room, a procedural improvement that forced subjects to look down into the pit. They then plucked from the air two other blocks floating in the Pit Room and dropped each on the same-colored Xs on the floor of the Living Room. The X-targets and green and blue blocks are visible in Figure 1. In this study, we investigated the effect of frame rate on presence. We hypothesized that the higher the frame rate, the greater the presence evoked.

Statistical significance. In this dissertation, we defined statistical significance at the 5% level. This is stated as $P < 0.050$. Findings significant at the 5% level are discussed as "demonstrated" or "shown". We also chose a method of statistical model construction in which we added variables to the model that were significant up to the 10% level ($0.05 \leq P < 0.10$).

Summary. Below we discuss our findings for the reliability, validity, multi-level sensitivity, and objectivity of the three physiological measures. We found that ΔHeart Rate bestmet our requirements for a measure of presence.

3 Physiological measures of presence

3.1 Reliability

Reliability is "the extent to which the same test applied on different occasions … yields the same result" [Sutherland, 1996]. Specifically, we wanted to know whether the virtual environment would consistently evoke a physiological reaction as the subject entered and remained in the Pit Room. The VE consisted of three rooms. Users started in the Training Room, which looked like a foyer or other small room in a house. They later moved to the Pit Room where, to get to the other side, they could either walk around the 20-foot drop to the room below using a two-foot-wide wooden catwalk or walk straight across – walking as if on a glass floor.

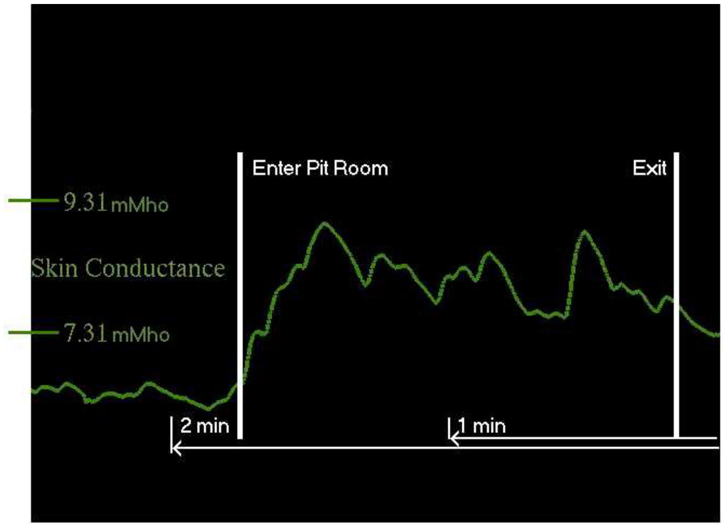


Figure 5. A typical skin conductance reaction to the Pit Room.

As we hypothesized, there were indeed significant physiological reactions to the Pit Room: heart rate and skin conductance were significantly higher and skin temperature was significantly lower in the Pit Room in all three studies. Figure 5 shows a typical skin conductance reaction to the Pit Room. Heart rate was higher in the Pit Room for 90% of the exposures to the VE, skin conductance was higher for nearly 95%, and skin temperature was lower for 90%. See Table 1.

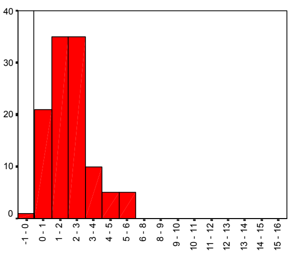
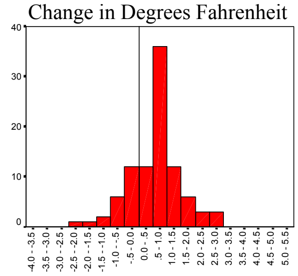
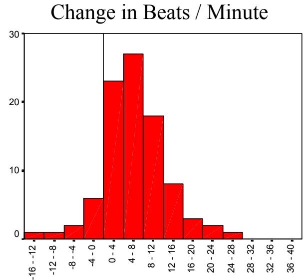
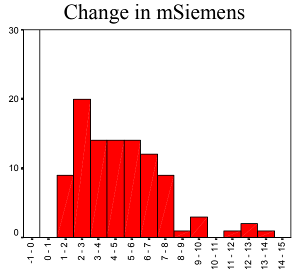
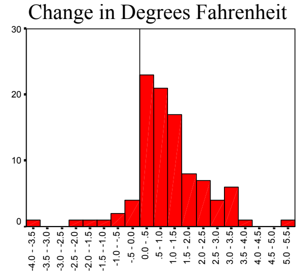
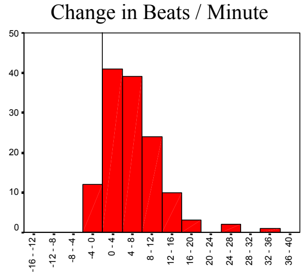
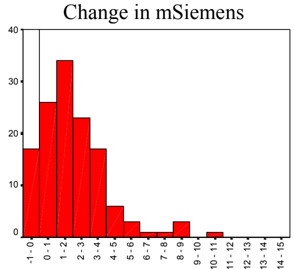
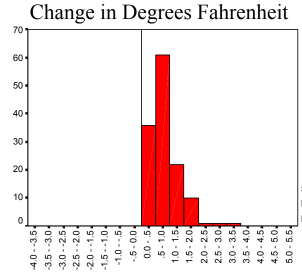
	Distribution of ΔHeart Rate	Distribution of ΔSkin Conductance	Distribution of ΔSkin Temperature
Effect of Multiple Exposures on Presence	**Not Available**	Change in mSiemens  Mean = 2.3 mSiemens Standard Deviation = 1.3 N = 112 Training Room Ave. = 8.3 mSiemens Count (Vals < 0.0) = 1 / 112	Change in Degrees Fahrenheit  Mean = 0.6 °F Standard Deviation = 0.9 N = 94 Training Room Ave. = 80.4 °F Count (Vals < 0.0) = 22 / 94
Effect of Passive Haptics on Presence	Change in Beats / Minute  Mean = 6.3 BPM Standard Deviation = 6.3 N = 92 Training Room Ave. = 85.9 BPM Count (Vals < 0.0) / N = 10 / 92	Change in mSiemens  Mean = 4.8 mSiemens Standard Deviation = 2.5 N = 100 Training Room Ave. = 6.7 mSiemens Count (Vals < 0.0) / N = 0 / 100	Change in Degrees Fahrenheit  Mean = 1.1 °F Standard Deviation = 1.2 N = 100 Training Room Ave. = 84.9 °F Count (Vals < 0.0) = 10 / 98
Effect of Frame Rate on Presence	Change in Beats / Minute  Mean = 6.3 BPM Standard Deviation = 5.8 N = 132 Training Room Ave. = 82.7 BPM Count (Vals < 0.0) = 12 / 132	Change in mSiemens  Mean = 2.0 mSiemens Standard Deviation = 2.0 N = 132 Training Room Ave. = 5.0 mSiemens Count (Vals < 0.0) = 17 / 132	Change in Degrees Fahrenheit  Mean = 0.8 °F Standard Deviation = 0.5 N = 132 Training Room Ave. = 83.6 °F Count (Vals < 0.0) = 0 / 132

Table 1. Differences in physiological measures between the Training Room and the Pit Room. The means of the distributions were significantly greater than zero ($P < 0.001$ for all measures in all studies).

We also wanted to know whether the physiological reactions to the environment diminished over multiple exposures. Since our hypotheses relied on presence in the VE evoking a stress reaction over a multiple exposures (2-12 exposures), we wanted to know whether physiological reactions to the VE would drop to zero or become unusably small due to habituation. In fact, each measure did decrease with multiple exposures (not necessarily significantly) in all studies (ΔSkin Temperature, Reported Presence, Reported Behavioral Presence, Observed Behavioral Presence, and ΔHeart Rate) or in all but one study (ΔSkin Conductance). None decreased to zero, though, even after twelve exposures to the VE. Table 2 shows the significant order effects.

A decrease in physiological reaction over multiple exposures would not necessarily weaken validity, since the literature shows that habituation diminishes the stress reactions to real heights and other stressors [Abelson, 1989; Andreassi, 1995]. Since all measures, not just the physiological measures decrease over multiple exposures, the decreases may not be due to habituation. There may be, as Heeter hypothesized, a decrease in presence evoked by the environment as novelty wears off [Heeter, 1992].

Order Effects	**ΔHeart Rate (ΔBPM)**	**ΔSkin Conductance (ΔmSiemens)**	**ΔSkin Temperature (Δ°F)**	**Reported Presence (Count "high")**	**Reported Behavioral Presence (Count "high")**	**Observed Behavioral Presence (Count Behvs.)**
Multiple Exposures	N/A	-0.7 (1st)	**-0.9 (1st)**	-	**-0.7 (1st)**	**-0.9 (Sess)**
Passive Haptics	-	-	-	**-0.8 (1st)**	**-0.4 (1st)**	-
Frame Rate	**-1.0 (Task)**	**-0.8 (1st)**	**-0.3 (1st)**	-	**-0.2 (Task)**	**-0.8 (1st)**

Table 2. Significant order effects for each measure in each study.
"(1st)" indicates a decrease in a measure after the first exposure only. "(Sess)" indicates a decrease in the measure over subsequent sessions (days). "(Task)" indicates a decrease over tasks on the same day. There was an order effect for each measure in at least one study. N/A is "Not available".

Orienting Effect. In general, each measure decreased after the first exposure. Moreover, there was a significant decrease after the first exposure for each measure except ΔHeart Rate in at least one of the studies (see Table 2). For physiological responses, this is called an *orienting effect* – a higher physiological reaction when one sees something novel [Andreassi, 1995]. Though this term traditionally refers to physiological reactions, we will also use the term for observed behavioral and reported reactions to the novel stimuli.

We attempted, with only partial success, to overcome this orienting effect by exposing subjects to the environment once as part of their orientation to the experimental setup and before the data-gathering portion of the experiment. In the Passive Haptics and Frame Rate studies, subjects entered the VE for approximately two minutes and were shown both rooms before the experiment started. These pre-exposures reduced but did not eliminate the orienting effects.

3.2 Validity

Validity is "the extent to which a test or experiment genuinely measures what it purports to measure" [Sutherland, 1996]. Since the concept of presence is itself vague and debatable, the question is then: How well do physiological reactions correlate with more traditional measures of presence? We investigated their correlations with several such measures.

Reported Presence. Among the physiological measures, ΔHeart Rate correlated best with the Reported Presence. There was a significant correlation in the Frame Rate study (corr. = 0.265, P = 0.002) and a weak and non-significant positive correlation (corr. = 0.034, P = 0.743) in the Passive Haptics study. In the Multiple Exposures study, where ΔHeart Rate was not available, ΔSkin Conductance had the highest correlation with Reported Presence (corr. = 0.245, P = 0.009).

Reported Behavioral Presence. Both ΔHeart Rate and ΔSkin Conductance correlated well with the Reported Behavioral Presence. ΔHeart Rate had the highest correlation, and a significant one with Reported Behavioral Presence in the Frame Rate study (corr. = 0.192, P = 0.028), and there was a weak and non-significant positive correlation between the two (corr. = 0.004, P = 0.972) in the Passive Haptics study. In the Multiple Exposures study, where ΔHeart Rate was not measured, ΔSkin Conductance had the highest correlation with reported behavioral presence (corr. = 0.290, P = 0.002). ΔSkin Conductance also had a non-significant positive correlation with Reported Behavioral Presence in the Passive Haptics Study (corr. = 0.106, P = 0.280).

The correlations of the physiological measures with the reported measures give some support to their validities. The validity of ΔHeart Rate appears to be well established by its correlation with the well-established reported measures. There was also support for the validity of ΔSkin Conductance from its correlation with reported measures, though not as strong support as for ΔHeart Rate.

There was little support for the validity of ΔSkin Temperature. As noted by McMurray, the measure suffers two limitations: 1) skin temperature response is slow (can take on the order of minutes for full effect) and is affected by many factors (sympathetic activity, muscular activity, etc.), 2) the sensors for detecting temperature changes are slow and can take on the order of one to three minute to fully register a change in temperature [McMurray, 1999]. We believe that these two lags combined with the limited time of exposure to the Pit Room (on the order of one minute) do not allow for enough time for useful measurement of ΔSkin Temperature.

Observed Behavioral Presence. Observed Behavioral Presence consistently correlated well with Reported Behavioral Presence, but it had mixed correlations with Reported Presence.

Following hypothesized relationships. According to Singleton, the validation process includes "examining the theory underlying the concept being measured," and "The more evidence that supports the hypothesized relationships [between the measure and the underlying concept], the greater one's confidence that a particular operational definition is a valid measure of the concept" [Singleton, 1993]. We hypothesized that presence should increase with frame rate and with the inclusion of the 1.5-inch wooden ledge. As presented in the next section, our findings followed the hypotheses – which helps validate the physiological reactions as measures of presence.

3.3 Sensitivity and multi-level sensitivity

Sensitivity is "the likelihood that an effect, if present, will be detected" [Lipsey, 1998]. The fact that the physiological measures reliably distinguished between subjects being in the Pit Room versus the Training Room assured us of at least a minimal sensitivity. All measures did so in every study. For example, heart rate increased an average of 6.3 beats / minute (BPM) in the Pit Room ($P < 0.001$) compared to the Training Room in both the Passive Haptics and Frame Rate studies. See Table 1 for full details of the means and standard deviations for each measure.

Multi-level sensitivity. For guiding VE technological development and for better understanding the psychological phenomena of VEs, we need a measure that reliably yields a higher value as a VE is improved along some goodness dimension, i.e., is *sensitive* to multiple condition values. We call this *multi-level sensitivity*. The Passive Haptics study provided us some evidence of multi-level sensitivity. Anecdotally, we have observed that walking into the Pit Room causes a strong reaction in users and this reaction is greater in "magnitude" than the differences in reaction to the Pit Room between any two experimental conditions (e.g. with and without the 1.5-inch wooden ledge). Therefore, we expected the differences among the conditions to be less than the difference between the two rooms. For example, in Passive Haptics, we expected there to be a significant difference in the physiological measures between the two conditions (with and without the 1.5-inch wooden ledge), but expected it to be less than the difference between the Training Room and Pit Room in the "lower" presence condition (without the 1.5-inch wooden ledge). For ΔHeart Rate, we did find a significant difference between the two conditions of 2.7 BPM (P = 0.016), and it was less than the inter-room difference for the without-ledge condition: 4.9 BPM. See Figure 6.

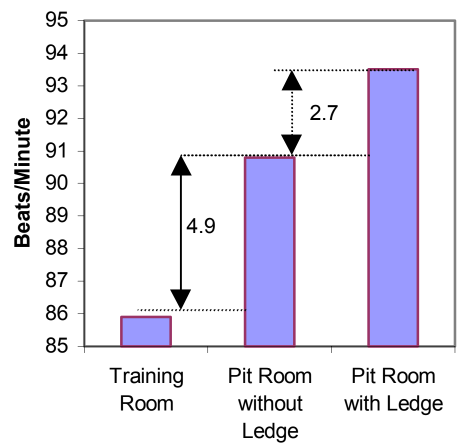


Figure 6. Heart Rate in the Passive Haptics study.

In the Passive Haptics study, we further tested the multi-level sensitivity by testing whether presence was significantly higher with the 1.5-inch wooden ledge. Presence as measured by ΔHeart Rate (2.7 BPM; P = 0.016), ΔSkin Conductance (0.8 mSiemens; P = 0.040), and Reported Behavioral Presence (0.5 "high" responses; P = 0.004) were significantly higher with the wooden ledge. Reported Presence had a strong trend in the same direction (0.5 "high" responses; P = 0.060). ΔSkin Temperature varied in the opposite direction; skin temperature decreased less when the 1.5-inch wooden ledge was present.

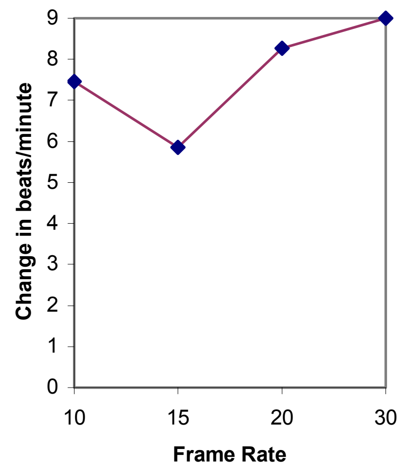


Figure 7. Average change in heart rate, after correcting for Loss of Balance, between Training Room and Pit Room at 10, 15, 20, and 30 frames per second.

In the Frame Rate study, we hypothesized that for graphics frame rates of 10, 15, 20, and 30 FPS, physiological reactions would increase monotonically with frame rate. They did not do exactly that (see Figure 7). During the 10 FPS condition, there was an anomalous reaction for all of the physiological measures and both the behavioral measures: Reported Behavioral Presence and Observed Behavioral Presence. That is, at 10 FPS, subjects had higher physiological reaction, reported more behavioral presence, and acted more present in the Pit Room. We believe that this reaction was due to discomfort, added lag, and reduced temporal fidelity while they were in the ostensibly dangerous situation of walking next to a 20-foot pit.

We also observed that subjects often lost their balance while trying to inch to the edge of the wooden platform at this low frame rate (Loss of Balance). Controlling for these Loss of Balance incidents improved the significance of the statistical model for ΔHeart Rate (3.5 BPM higher when Loss of Balance; P = 0.014) and brought the patterns of responses closer to the hypothesized monotonic increase in presence with frame rate – but did not completely account for the increased physiological reaction at 10 FPS. Loss of Balance was not significant in any other model.

Beyond 10 FPS, ΔHeart Rate performed well after statistically controlling for Loss of Balance. ΔHeart Rate significantly increased (3.2 BPM; P = 0.004) between 15 FPS and 30 FPS and between 15 FPS and 20 FPS

(2.4 BPM; P = 0.024). There was also a non-significant increase between 20 FPS and 30 FPS (0.7 BPM; P = 0.483) and a non-significant decrease between 10 FPS and 15 FPS (1.6 BPM; P = 0.134). Reported Presence, Reported Behavioral Presence, and ΔSkin Temperature also increased with frame rate from 15-20-30 FPS, but with less distinguishing power. These findings support the multi-level sensitivity of ΔHeart Rate but do not support those of ΔSkin Conductance or ΔSkin Temperature.

3.4 Objectivity

The measure properties of reliability, validity, and multi-level sensitivity are established quantitatively. Objectivity can only be argued logically. We argue that physiological measures are inherently better shielded from both subject bias and experimenter bias than are either reported measures or measures based on behavior observations. Reported measures are liable to subject bias – the subject reporting what he believes the experimenter wants. Post-experiment questionnaires are also vulnerable to inaccurate recollection and to modification of impressions garnered early in a run by impressions from later. Subject reporting during the session, whether by voice report or by hand-held instrument, intrudes on the very presence illusion one is trying to measure.

Observed measures are obviously liable to scorer bias. The use of videotape helps by making it possible for the scorer to do replays and to make considered judgments. In some studies, however, it is impossible to hide the condition from the experimenter scoring the measure. For example, in the Passive Haptics study, the experimenter can see the 1.5-inch wooden ledge on the videotape. Behavioral measures are also somewhat liable to subject bias. We observed occasional intentionally exaggerated or intentionally suppressed fear behaviors near the pit.

Physiological measures, on the other hand, are much harder for subjects to affect, especially with no bio-feedback. They are not liable to experimenter bias, if instructions given to the users are properly limited and uniform. We read instructions from a script in the Multiple Exposures study. We improved our procedure in the later Passive Haptics and Frame Rate studies by playing instructions from a compact disk player located in the real laboratory and represented with a virtual radio in the VE.

3.5 Summary and discussion

The data presented here show that physiological reactions can be used as reliable, valid, multi-level sensitive, and objective measures of presence. These findings are summarized in Table 3. Of the physiological measures, ΔHeart Rate performed the best. It significantly differentiated between the Training Room and the Pit Room and this reaction faded over multiple exposures, but never to zero. It correlated with the well-established reported measures. It distinguished between the inclusion of passive haptics and among frame rates after 10 FPS. As we argued above, it is objective. In total, it satisfies all of the requirements for a reliable, valid, multi-level sensitive, and objective measure of presence.

	Reliable	Valid	Sensitive	Objective
ΔHeart Rate	✓	✓	✓	✓
ΔSkin Conductance	✓	✓/✗	✓/✗	✓
ΔSkin Temperature	✓	✗	✗	✓

Table 3. Summary of findings.

ΔSkin Conductance and ΔSkin Temperature have some, but not all, of the properties we desire in a measure of presence. In our investigation, at least one of the properties of reliability, validity, multi-level sensitivity, and objectivity were not met for each of these measures. We believe that ΔSkin Temperature performed poorly because the exposure to the VE was too short for useful measurement of changes in skin temperature. We do not have a theory as to why ΔSkin Conductance was not multi-level sensitive in the Frame Rate study.

We found that ΔHeart Rate satisfied the requirements for a presence measure for our VE, which evokes a strong reaction, but it may not for all VEs. If this did not work for so strong a stimulus, it would not work for less stressful VEs. Our investigation is only a first step. More investigation would be needed to determine if physiological reaction could also work as a measure of presence for less stressful or non-stressing VEs. A group at Goldsmith's college is currently looking into whether physiological reaction can be used as a measurement of presence in stressful and relaxing 3D TV presentations [Dillon, 2001].

Another desirable aspect of a measure is ease of operationalization. We did not measure the time taken for each measure, but after running many subjects we can say with some confidence that use of the physiological monitoring and the presence questionnaire added approximately the same amount of time to the experiment, with the questionnaire taking a little less time. The observed behavioral measure did not add to the time to conduct the experiment, but added considerable time after the experiment. It took about five minutes of both the experimenter's and subject's time to put the physiological sensors on and take them off for each exposure. It took about an extra minute at the beginning and end of each set of exposures to put on and take off the ECG sensor – it was left on between exposures on the same day. It took subjects about five minutes to fill out the 16-item UCL Presence Questionnaire. Time to review the videotape of each session was the longest – this took an average of eleven minutes per exposure (around 25 hours to review 132 sessions in the Frame Rate experiment, the only one for which we kept track of the time spent). A benefit of the presence questionnaire was that the experimenter was free to do something else while the subjects filled them out. The observed behavioral measure took no additional time for the subject and very little training was needed to do the scoring. It took some training to learn the proper placement of the physiological equipment on the hands and chest of the subject – thirty minutes would probably be sufficient for most graduate students. We avoided loosing time while connecting the physiological monitoring equipment by having one experimenter start up the VE while the other connected the physiological sensors.

Another aspect of ease of operationalization is ease of use. No subjects reported difficulties with the questionnaires. None had a problem with being video-recorded during the sessions. Technical problems, however, plagued our videotaped measure. Only about one in ten subjects reported noticing the physiological monitoring equipment on the hands during the VE exposures. Our experiment, though, was designed to use only the right hand, keeping the sensor-laden left hand free from necessary activity. No subjects reported noticing the ECG sensor once it was attached. In fact, many subjects reported forgetting about the ECG electrodes when prompted to take them off at the end of the day. There are groups investigating less cumbersome equipment including a physiological monitoring system that subjects wear like a shirt [Cowings, 2001].

Overall, questionnaires and physiological monitoring were easy to operationalize, the observed behavioral measure was less so.

4 Physiological reactions as between-subjects measures

We conducted all of the studies as within-subjects to avoid the variance due to natural human differences. That is, subjects experienced all of the conditions for the study in which they participated. This allowed us to look at relative differences in subject reaction among conditions and to overcome the differences among subjects in reporting (for the questionnaires), physiological reaction, and behavior (for the Observed Behavioral Presence measure).

The UCL questionnaire has been used successfully between-subjects [Usoh, 1999]. We suspected, however, that physiological reaction would not perform as well if taken between-subjects. Specifically, we expected that between-subjects physiological measures would not be able to significantly differentiate among presence conditions, since the variance among subjects would mask, at least in part, the differences in physiological reaction evoked by the different conditions. We also expected correlations with the reported measures to be reduced, since individual differences in physiological reaction and reporting would confound the correlations. For example, consider a subject presented with a low presence condition who reported low presence. If the subject was highly physiologically reactive, however, he still could have had a high physiological reaction relative to the group average – even if it were lower than it would have been if he was presented with a high presence condition. Such a case would reduce the correlations.

We expected that there would still be a consistent physiological reaction to the Pit Room, since we expected such a reaction for every exposure to the VE. We expected the significance to be slightly lower, however, because of the reduced size of the data set.

We investigated these hypotheses in this section by analyzing the data using *only the first task* for each subject – eliminating order effects and treating the reduced data sets as between-subjects experiments. That is, we treat each experiment as if only the first task for each subject was run. This means that the analysis uses only 10 data points (10 subjects – first exposure only) for the Multiple Exposures study, 52 data points for the Passive Haptics study, and 33 data points for the Frame Rate study.

Reliability between-subjects: Physiological reaction in the Pit Room. As suspected, all of the physiological reactions were significantly higher in the Pit Room when analyzing between-subjects. See Table 4. Also, as expected, the significance values for the differences were lower when looking at only the first exposure than when looking at all exposures, due to the reduced number of data points. The means for the differences in physiological reactions for the first exposures were *higher* than for the full data set (except for ΔSkin Conductance in the Passive Haptics study). This follows, since we observed physiological orienting effects. The fraction of exposures in which heart rate and skin conductance were higher and skin temperature was lower in the Pit Room as compared to the Training Room is the same or higher than these for the within-subjects data, and all are 90% or better. Heart Rate in Passive Haptics is the only exception (85% vs. 89%). See Table 4.

		First Exposure Only				All exposures			
Study	Variable	**Mean**	***P***	N	**Count < 0**	**Mean**	***P***	N	**Count < 0**
Multiple Exposures	**ΔSkin Conductance**	**2.9**	*.002*	9	0	2.3	*.000*	112	1
	ΔSkin Temperature	**1.2**	*.015*	7	0	0.6	*.000*	94	22
Passive Haptics	**ΔHeart Rate**	**6.2**	*.000*	46	7	6.3	*.000*	92	10
	ΔSkin Conductance	4.7	*.000*	50	0	**4.8**	*.000*	100	0
	ΔSkin Temperature	**1.1**	*.000*	49	3	1.1	*.000*	98	10
Frame Rate	**ΔHeart Rate**	**8.1**	*.000*	33	3	6.3	*.000*	132	12
	ΔSkin Conductance	**2.6**	*.000*	33	1	2.0	*.000*	132	17
	ΔSkin Temperature	**1.0**	*.000*	33	0	0.8	*.000*	132	0

Table 4. Means and significance for one-sample t-test (test to see if the mean was significantly higher than zero). Also shown is the count of times that the measure was below zero. Physiological reaction, shown here for first task only and for all tasks together, was higher in the Pit Room. The higher mean is shown in bold face.

Validity between-subjects: Correlation with established measures. As expected, physiological reactions did not correlate as well with the questionnaires when analyzing between-subjects. There were no significant correlations between the physiological measures and any of the questionnaire measures in any of the studies.

Multi-level sensitivity between-subjects: Differentiating among presence conditions. We expected inter-subject variation in physiological reaction to mask the differences in physiological reactions evoked by the presence conditions (e.g. var[illegible] frame rates). Contrary to this expectation, though, the physiological measures **did** differentiate among the conditions: physiological reaction to the Pit Room was significantly higher that to the Training Room for all measures in all studies (described above), and we found significant differences in the physiological measures among conditions in both the Passive Haptics and Frame Rate studies. (The condition was not varied in the Multiple Exposures study.)

In the Passive Haptics study, both ΔHeart Rate and ΔSkin Conductance performed well as between-subjects presence measures. In the base statistical model (not correcting for anything) both varied in the expected direction with some power (P = 0.097 for ΔHeart Rate; P = 0.137 for ΔSkin Conductance). After correcting for subjects' level of computer game playing, the significance for ΔHeart Rate was reduced (P = 0.180).

In the Frame Rate study, ΔHeart Rate performed well, but ΔSkin Conductance did not follow hypothesized patterns. ΔHeart Rate had an anomalous physiological reaction at 10 FPS. This was also the case for ΔHeart Rate in the full data set (compare Figures 7 and 8). Additionally, ΔHeart Rate differentiated among presence conditions: ΔHeart Rate at 30 FPS was higher than at 15 FPS and this difference was nearly significant (7.2 BPM; P = 0.054).

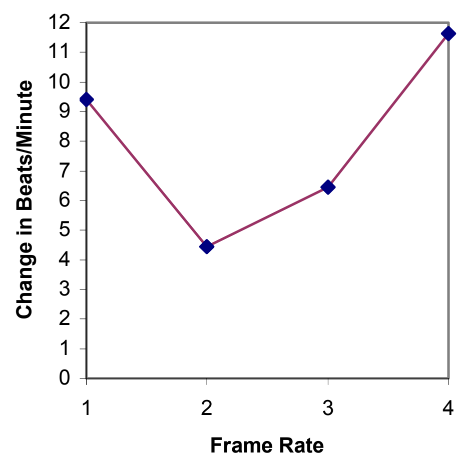


Figure 8. Between-subjects analysis: Response graph of ΔHeart Rate.

Overall, ΔHeart Rate shows promise as a between-subjects measure of presence. Though it did not correlate well with the reported measures (between-subjects), it did differentiate among the conditions with some statistical power in Passive Haptics and Frame Rate. ΔSkin Conductance and ΔSkin Temperature did not show as much promise as between-subjects measures. For more discussion of physiological reaction as between-subjects measures of presence, see Section 4.4.1.

5 The results of the studies

To establish the properties of the physiological presence measures, we conducted three controlled user studies, plus the pilot studies necessary to debug environment, equipment, and procedures. Each was a dual-purpose study, both contributing to the development and testing of physiological presence measures (discussed above) and investigating some other aspect of VEs. Below, we summarize the results and findings of the three studies individually.

5.1 Effect of multiple exposures on presence (Multiple Exposures)

Hypothesis: The presence-inducing power of a VE declines with multiple exposures, but not to zero.

Ten subjects had three exposures to the VE each day on four separate days – a total of twelve exposures each. The VE and the task were the same for each exposure.

Results. There was a significant order effect for each measure in at least one of the three studies, and when not significant, the trends were in the same direction. Table 2 summarizes the order-effect results significant at the $P < 0.050$ level (**bold**) and $P < 0.100$ (normal text), not only for this study, but also for the two subsequent ones. The existence and magnitude of the significant order effects in all the measures supports the hypothesis that all presence measures decreases over 12 exposures to the same VE, but not to zero.

5.2 Effect of passive haptics on presence (Passive Haptics)

Hypothesis: Supplementing a visual-aural VE with even rudimentary, low-fidelity passive haptics cues significantly increases presence.

This experiment was only one of a set investigating the passive haptics hypothesis. The detailed design, results, and discussion for the set are reported elsewhere [Insko, 2001].

Design. Fifty-two subjects each had two exposures on separate days. For the passive haptics condition, the virtual ledge in the Pit Room was augmented with a registered real plywood ledge, 1.5 inches high. A user in a HMD, unable to see the real world, could feel the edge of the ledge with the foot. The 1.5-inch height was selected so that the edge-probing foot did not normally contact the real laboratory floor where the virtual pit was. Each subject experienced the environment with the 1.5-inch wooden ledge (“high” presence condition) and without it ("low" presence condition). Presentation of the conditions was counterbalanced across subjects. Figure 9 shows a subject standing on the 1.5-inch wooden ledge.

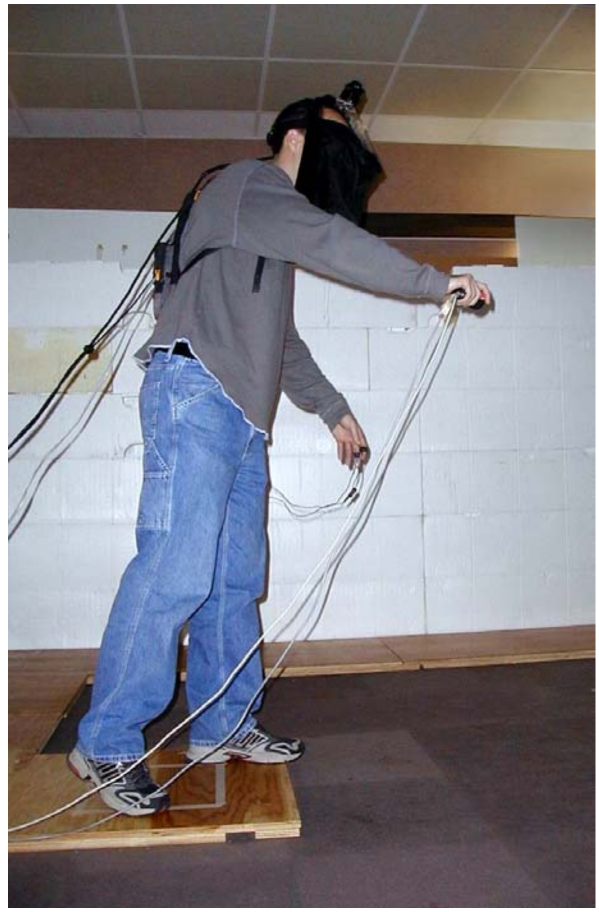


Figure 9. A subject drops a block into the virtual pit. He is standing on the edge of the 1.5-inch wooden ledge. The physiological monitoring equipment is attached to his left hand.

Results. As discussed above these measures were significantly higher with the 1.5-inch wooden ledge: ΔHeart Rate was significantly higher ($P = 0.016$) with the wooden ledge than without it. Reported Behavioral Presence ($P = 0.004$) and ΔSkin Conductance ($P = 0.040$) were also significantly higher. Reported Presence had a strong trend ($P = 0.060$) in the same direction.

5.3 Effect of frame rate on presence (Frame Rate)

Hypothesis: As frame rate increases from 10, 15, 20, 30 frames/second, presence increases.

Thirty-three subjects each had four exposures to the same VE and task, at each of several frame rates. Presentation order was counterbalanced across subjects.

Results. Discussed in Section 1.4 above, the hypothesis was confirmed for 15, 20, 30, FPS, but 10 FPS gave anomalous results.

6 Future work

Given a compelling VE and a good presence measure, the obvious strategy would be to degrade the VE quality parameters one at a time so as to answer: What makes a VE compelling? What are the combinations of minimum system characteristics to achieve this?

For example, we hope to study

- Aural localization,
- Visual Detail,
- Lighting realism,
- Self-avatar fidelity,
- Realistic physics in interactions with objects, and
- Interactions with other people or agents.

We want to begin to establish trade-offs for presence evoked: Is it more important to have lag below 50 ms or frame rate above 20 FPS? These tradeoffs could eventually lead to identification of *isosurfaces* for presence, as described by Ellis [Ellis, 1996]. In particular, physiological reaction satisfies his requirement that "an independent measure of human performance or some other independent characteristic of the virtual environment should be shown to be determined by equivalence classes." We could compare the effect of varying system parameters (e.g. lag, frame rate, visual realism, etc.) on the extent of physiological reaction evoked. Assuming repeatable results, sets of VE system parameter (e.g. *Lag* = Lag_j, *Frame Rate*=FR_j, *Use of localized sound* = Yes, etc.) that evoke equal physiological reactions would be in the same equivalence class.

Future work should include using VEs that evoke more subtle reactions and different reactions than ours. To use physiological reaction as a measure of presence, *feeling presence* in the VE must evoke a physiological reaction distinct from that of the laboratory environment. Ours did this by evoking in subjects the perceived danger of moving about near a height. Other VEs might evoke presence in other stressful environments, exciting environments, interesting environments, relaxing environments, etc. If future investigators are unsure of the physiological reaction that the VE should evoke, then a controlled study in the corresponding real environment should be conducted.

Further investigation should also include VEs that expose the subject to the stressor for less time. In a study with a shorter exposure to the stressor, one should consider three things:

1) There is a lag of 2-3 seconds from time of stimulus (exposure to stressor) to onset of reaction for all physiological reaction including heart rate, skin conductance, and skin temperature [Andreassi, 1995].

2) Heart rate is affected by respiration. For accurate measurement, heart rate should be averaged over one or two respiration cycles. A respiration cycle averages 4 seconds [Seidel, 1995].

3) Skin temperature does not work well for exposures of less than 2 minutes.

We must eliminate the cables that tether subjects to the monitoring, tracking, and rendering equipment. Our subjects reported this encumbrance as the greatest cause of breaks in presence.

7 Bibliography

Abelson, J. L. and G. C. Curtis (1989). Cardiac and neuroendocrine responses to exposure therapy in height phobics. *Behavior Research and Therapy,* 27(5): 561-567.

Andreassi, J. L. (1995). *Psychophysiology: human behavior and physiological response*. Hillsdale, N.J., Lawrence Erlbaum Associates.

Cobb, S., S. Nichols, A. Ramsey and J. R. Wilson (1999). Virtual reality induced symptoms and effects (VRISE). *Presence: Teleoperators and Virtual Environments,* 8(2): 169-186.

Cohen, D. C. (1977). Comparison of self-report and overt-behavior procedure for assessing acrophobia. *Behavior Therapy,* 8: 17-23.

Cowings, P., S. Jensen, D. Bergner and W. Toscano (2001). A lightweight ambulatory physiological monitoring system. Moffett Field, California, Ames Research Center.

Dihn, H. Q., N. Walker, L. F. Hodges, C. Song and A. Kobayashi (1999). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. Proceedings of *Virtual Reality*, 222-228, Houston, Texas, USA.

Cath Dillon, Edmund Keogh, Jonathan Freeman, and Jules Davidoff. Presence: Is Your Heart In It? 3rd Int. Wkshp. on Presence

Eberhart, R. C. and P. N. Kizakevich (1993). Determining physiological effects of using VR equipment. Proceedings of *First Annual Conference, Virtual Reality and Persons with Disabilities*, 47-49.

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(2): 247-259.

Emmelkamp, P. and M. Felten (1985). The process of exposure *in vivo*: cognitive and physiological changes during treatment of acrophobia. *Behavior Research and Therapy,* 23(2): 219.

Fontaine, G. (1992). The experience of a sense of presence in intercultural and international encounters. *Presence: Teleoperators and Virtual Environments,* 1(4): 482-490.

Guyton, A. C. (1986). Basic characteristics of the sympathetic and parasympathetic function. In *Textbook of Medical Physiology*, 688-697. Philadelphia, W.B. Saunders Company.

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

Held, R. M. and N. I. Durlach (1992). Telepresence. *Presence: Teleoperators and Virtual Environments,* 1(Winter): 109-112.

Hendrix, C. and W. Barfield (1996). The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments,* 5(3): 290-301.

Hodges, L., B. Rothbaum, R. Kooper, D. Opdyke, J. Willford, T. Meyer and M. North (1994). Presence as the defining factor in a VR application. Technical Report GVU-94-06. Georgia Tech University, Graphics, Visualization, and Usability Center.

Hodges, L. F., B. O. Rothbaum, R. Kooper, D. Opdyke, T. Meyer, M. North, J. J. de Graaff and J. Williford (1995). Virtual environments for treating the fear of heights. *IEEE Computer,* 28(7): 27-34.

Insko, B. (2001). Passive haptics significantly enhance virtual environments. Doctoral Dissertation. Computer Science Department. Chapel Hill, NC. University of North Carolina.

Jones, O. R. and D. A. Grant (1948). Category difficulty study on the University of Wisconsin Card Sorting Test. *American Psychologist,* 3.

Jorgensen, C., J. Ogden, K. Willis, M. Blessing, K. A. Caudell, G. Patrick and T. P. Caudell (1997). Locomotion in a virtual environment: Performance measures and physiological responses. Proceedings of *Human Factors and Ergonomics Society 41st Annual Meeting*, 1148-1151, Albuquerque, New Mexico.

Kennedy, R., N. Lane, K. Berbaum and M. Lilienthal (1993). A simulator sickness questionnaire(SSQ): A new method for quantifying simulator sickness. *International Journal of Aviation Psychology,* 3(3): 203-220.

Kleinbaum, D., L. Kupper, K. Muller and A. Nizam (1998). *Applied regression analysis and other multivariate methods*.

Lipsey, M. W. (1998). Design sensitivity: Statistical power for applied experimental research. In L. Brickman and D. J. Rog, Eds., *Handbook of applied social research methods*, 39-68. Thousand Oaks, California, Sage Publications, Inc.

Lombard, M. and T. Ditton (1997). At the heart of It all: The concept of presence. *Journal of Computer Mediated Communication,* 3(2).

Lombard, M. and T. Ditton (1999). Across media: Toward a standardized conceptual and operational definition of presence. 2nd Int. Wkshp. on Presence.

McMurray, D. R. (1999). Director of Applied Physiology lab, University of North Carolina. Personal Communication.

Meehan, M. (2000a). Measuring presence objectively. *British Psychological Society Work Shop on Presence and Its Measures*.

Meehan, M. (2000b). An objective surrogate for presence: Physiological response. *3rd Int. Wkshp. on Presence*.

Meehan, M., L. Pugnetti, F. Riva, E. Barbiera, L. Mendozzi and C. E. (2000c). Peripheral responses to a mental-stress inducing virtual environment experience. Proceedings of *International Conference on Disability, Virtual Reality and Associated Technologies*, Sardinia, Italy.

Mine, M. R. (1993). Characterization of end-to-end delays in head-mounted display systems. Technical Report TR93-001. Chapel Hill, NC, University of North Carolina.

Plato (380 B.C.). Book 5. In *The Republic*. Indianapolis, Hacket Publishing Company.

Psotka, J. M., S. Davison and S. A. Lewis (1993). Exploring immersion in virtual space. *Virtual Reality Systems,* 1(2): 70-82.

Pugnetti, L., L. Mendozzi, A. M. Cattaneo, S. Guzzetti, C. Mezzetti, C. Coglati, D. A. E., A. Brancotti and D. Venanzi (1995). Psychophysiological activity during virtual reality (VR) testing. *Journal of Psychophysiology,* 8(4): 361-362.

Pugnetti, L., L. Mendozzi, E. Barbieri, F. D. Rose and E. A. Attree (1996). Nervous system correlates of virtual reality experience. Proceedings of *1st European Conference on Disability, Virtual Reality and Associated Technologies*, 239-246.

Pugnetti, L., M. Meehan, L. Mendozzi, F. Riva, E. Barbiera and C. E. (2000). More on Central Nervous System Correlates of Virtual Reality Testing. Proceedings of *International Conference on Disability, Virtual Reality and Associated Technologies*, Sardinia, Italy.

Regenbrecht, H. T. and T. W. Schubert (1997). Measuring presence in virtual environments. *Human Computer Interface International*, San Francisco, USA.

Sallnas, E. L. (1999). Presence in multimodal devices. 2nd Int. Wkshp. on Presence.

SAS (1990). *SAS/ STAT User's Guide, Version 6, Fourth Edition*. Cary, NC, USA, SAS Institute Inc.

Schloerb, D. W. (1995). A quantitative measure of telepresence. *Presence: Teleoperators and Virtual Environments,* 4(1): 64-80.

Seidel, H. and H. Herzel (1995). Modeling heart rate variability due to respiration and baroreflex. In E. Mosekilde and O. G. Mouritsen, Eds., *Modeling the dynamics of biological systems*, Springer Verlag.

Sheridan (1992). Musings on telepresence and virtual presence, telepresence. *Presence: Teleoperators and Virtual Environments,* 1(Winter): 120-126.

Sheridan, T. B. (1996). Further musings on the psychophysics of presence. *Presence: Teleoperators and Virtual Environments,* 5(2): 241-246.

Short, J., E. Williams and B. Christie (1976). *The social psychology of telecommunications*. London, Wiley.

Singleton, R. A., B. C. Straits and M. M. Straits (1993). *Approaches to Social Research*. New York, Oxford University Press.

Slater, M. and M. Usoh (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments,* 2(3): 221-233.

Slater, M., M. Usoh and A. Steed (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments,* 3(2): 130-144.

Slater, M., M. Usoh and Y. Chrysanthou (1995a). The influence of dynamic shadows on presence in immersive virtual environments. *Virtual Environments '95*, 8-21.

Slater, M., M. Usoh and A. Steed (1995b). Taking steps: The influence of a walking technique on presence in virtual reality. Proceedings of *ACM Transactions on Computer Human Interaction (TOCHI)* 2(3): 201-219.

Slater, M., V. Linakis, M. Usoh and R. Kooper (1996). Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess. Proceedings of *ACM Virtual Reality Software and Technology (VRST)*, 163-172.

Slater, M. (1999). Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments,* 8(5): 560-565.

Slater, M. and A. Steed (2000). A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments,* 9(5): 413-434.

SPSS (1999). SPSS Release 10.0.1 Standard Version, SPSS.

Steed, A., M. Slater, A. Sadagic, A. Bullock and J. Tromp (1999). Leadership and collaboration in shared virtual environments. Proceedings of *Virtual Reality*, 112-115, Houston, Texas.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication,* 42(4): 73-93.

Sutherland, S. (1996). *The international dictionary of psychology*. New York, The Crossroads Publishing Company.

Thie, S. and J. Van Wijk (1998). A general theory on presence. *1st Int. Wkshp. on Presence*.

Uno, S. and M. Slater (1997). The sensitivity of presence to collision response. Proceedings of *Virtual Reality Annual International Symposium (VRAIS)*.

Usoh, M., K. Arthur, M. Whitton, R. Bastos, A. Steed, M. Slater and F. Brooks (1999). Walking > walking-in-place > flying in virtual environments. Proceedings of ACM *SIGGRAPH 99*.

Usoh, M., E. Catena, S. Arman and M. Slater (2000). Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments,* 9(5): 497-503.

Ward, M., R. Azuma, R. Bennett, S. Gottschalk and H. Fuchs (1992). A demonstrated optical tracker with scalable work area for head mounted displays. *Proc. of the 1992 Symposium on Interactive 3D Graphics,* 25(2): 43-52.

Webster (1994). *Webster's New World Dictionary*.

Weiderhold, B. K., R. Gervirtz and M. D. Weiderhold (1998). Fear of flying: A case report using virtual reality therapy with physiological monitoring. *CyberPsychology and Behavior,* 1(2): 97-104.

Welch, G. and G. Bishop (1997). SCAAT: Incremental tracking with incomplete information. Proceedings of ACM *SIGGRAPH 97*, 333-344.

Witmer, B. G. and M. J. Singer (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments,* 7(3): 225-240.

Yamaguchi, T. (1999). Physiological studies of human fatigue by a virtual reality system. *Presence: Teleoperators and Virtual Environments,* 8(1): 112-124.