Visual Realism Enhances Realistic Response in an Immersive Virtual Environment

Mel Slater, Pankaj Khanna, Jesper Mortensen, and Insu Yu ▪ *University College London*

A study comparing real-time recursive ray tracing with ray casting in an immersive visual environment shows how greater visual realism induces greater participant presence.

Participants in an immersive virtual environment interact with the scene from an egocentric point of view—that is, where there bodies appear to be located—rather than from outside as if looking through a window. People interact through normal body movements, such as head-turning, reaching, and bending, and—within the tracking limitations—move through the environment or effect changes within it in natural ways.

A major application of such systems is rehearsal or training for situations too dangerous or impractical to carry out in real life, such as vehicle simulators. Another major application is in psychotherapy, where patients might experience anxiety-provoking events with the knowledge that nothing real is happening but their normal anxiety responses are still induced and ultimately reduced through repeated exposure. In these scenarios, the application's success relies on participants responding to situations within the virtual environment as if these were real. If not, they couldn't transfer what they learned to the real world. Such "response as if real" provides an operational definition of the concept of *presence*,[1] where response is considered at multiple levels: subjective, behavioral, and physiological (such as changes in heart rate).

In this article, we consider the impact of *visual realism* on presence. Visual realism has two components—geometric realism (the virtual object looks like the real object) and illumination realism (referring to the fidelity of the lighting model). Both types have static and dynamic aspects. An object might look statically realistic, but over time its dynamic changes may not. For example, an object representing a human might look realistic but not behave realistically. Similarly, good static lighting might be achieved with a method such as radiosity, which precomputes all diffuse interreflections in the environment, but shadows might not move in conjunction with the corresponding objects. Here we focus on the impact of illumination, such as that achieved by real-time ray tracing.

One hypothesis says that visual realism isn't important for presence. This is based on the idea that the human perceptual system works in a top-down manner, building apparently complete environmental representations from a few minimal cues.[1] For example, in principle we should respond appropriately even in the case of wire-frame rendering—provided that there is high frame rate, wide field of view, low latency, stereo, and a head-tracked system. Indeed, some have argued that should the level of displayed realism improve "too much," this could lead to degradation in response, because human observers would be more likely to notice small imperfections. This is Masahiro Mori's 1970 "Uncanny Valley" (UV) hypothesis—that improvements in quality might result in improvements in response up to a point after which there might be a sudden dip in response due to defect magnification.[2]

However, the UV hypothesis is convenient where, at least in the context of real-time rendering for immersive virtual environments (VEs), it hasn't been feasible until recently to produce high-quality visual rendered environments. Moreover, evidence hasn't been clear regarding the impact of

 Published by the IEEE Computer Society

visual quality on presence. In this article, we show that when recursive real-time ray tracing is used to render an immersive VE, participants experience significantly higher appropriate anxiety than when the same environment is rendered with ray casting (ray tracing, but with single eye-to-object rays). We used an environment that displays a precipice, a pit that the participant looks over—an environment frequently used in presence evaluations.

Background

Early papers that addressed the impact of visual realism on presence compared environments which have different levels of visual detail and pictorial realism. Claudia Hendrix and Woodrow Barfield[3] describe the fact that many researchers have found greater *reported presence* with higher visual realism in both senses. By reported presence we mean the sense of being in the virtual environment as assessed by questionnaires. Another study compared reported and behavioral presence across rendering styles that didn't support shadows, supported static shadows, or supported dynamically changing shadows. Mel Slater and his colleagues found greater behavioral and reported presence for static shadows compared to no shadows and for dynamic shadows compared to static shadows.[4] The dynamic shadows were computed using an adaptation of the Binary Space Shadow Volume Binary Space Partition algorithm.

In a questionnaire-based study using two levels of radiosity and flat shading, Katerina Mania and Andrew Robinson found no difference in reported presence among the three conditions.[5] Finally, in an experiment set in the "pit room" environment, similar to the one used in our experiment, the scene was displayed at various levels of illumination realism (wire frame, without and with textures, and with radiosity).[6] In that experiment, Paul Zimmons and Abigail Panter recorded physiological measures and questionnaire responses. Subjects in every condition exhibited significantly increased heart rate when they encountered the pit, although there were no significant differences in heart rate or reported presence between different rendering conditions.

Results to date show no clear message regarding the impact of the visual realism level on reported or behavioral presence. Evidence supports conclusions that the type of rendering might or might not impact presence. Moreover, we believe many previous studies were flawed, because they were based on within-group designs (see the "Between-Groups vs. Within-Groups" sidebar) in which participants would experience all experimental conditions (different levels of visual realism) and therefore realize the purpose of the experiment. It's important to note that this doesn't apply to the study most directly comparable with the one reported here.[6]

Our study differs in three important ways from those reported previously. First, we use real-time recursive ray tracing. Although it's not full global illumination, it correctly simulates light transport between specular surfaces and therefore generates dynamically changing reflections and shadows for point light sources. In particular, it

Between-Groups vs. Within-Groups

A between-group design is where each participant experiences only one condition so that the comparisons are between the different groups rather than within individuals. Each group is drawn from the same population so that they are matched on the average. In addition, data can be recorded about their demographic situation (age, gender, status), and this information can further statistically equalize the groups in the later analysis. A between-group design has the advantage that participants don't learn the purposes of the experiment because they experience only one condition. It has the disadvantage that typically more people are needed than for within-group designs to achieve good statistical results. A within-group design is one in which each participant experiences every condition and the order of presentation is randomized across participants. This has the advantage that each person is compared to himself or herself, and fewer participants are needed than with a between-groups design to achieve good statistical results.

In general, the problem with within-groups designs in virtual reality experiments is that conditions are not symmetric. In other words, it's certain that the experience of one condition (for example, ray tracing) will affect the experience of the other condition (ray casting), and presenting the conditions in different orders makes no difference to this. A second problem is that the participants obviously realize, or might think they realize, the purpose of the experiment, because they see all conditions. This could also bias their responses, especially if they feel that they should give the kinds of answers that they believe would please the experimenters. Finally, there is a question of adaptation. Having experienced the precipice once, a participant's response might be different on the second exposure. In our earlier brief report of some aspects of this experiment's results,[1] we showed that responses were not symmetric comparing the between-group conditions only and the within-group conditions with regard to subjectively reported presence. In this article we report only on the between-group condition, from which the most reliable results can be obtained.

References

1. P. Khanna et al., "Presence in Response to Dynamic Visual Realism: A Preliminary Report of an Experiment Study," *Proc. ACM Symp. Virtual Reality Software and Technology*, ACM Press, 2006, pp. 364–367.

generates reflections and shadows of the virtual body that represent the participant immersed in the (head-mounted-display-generated) virtual reality. Although one study had dynamic shadows, these were limited to a single object on a flying trajectory and didn't include shadows from the participant's virtual body.[4] Second, although we use questionnaires to assess the subjective element of presence (the sense of "being there" in the place depicted by the VE), we also analyze the physiological responses of the participants. In particular, we carry out electrocardiogram (ECG) analysis to determine stress levels. Third, our study is between-groups, which means the results are based on each participant's experience of only one condition (recursive ray tracing or ray casting) and therefore can't be biased by participants' understanding the study's purposes.

The pit environment was inspired by the famous "visual cliff" experiments, which investigated depth perception in different animals, in particular human babies.

The Pit Room Environment

The scene in our experiment was a variation of the pit room (see Figure 1). In such scenarios the participant enters a virtual training room. This is typically an ordinary room with some furniture where the participant gets accustomed to the environment and learns to carry out tasks relating to the particular experiment. Then he or she is required to move into an adjoining room through an open door. This room seems normal at first, but then the participant sees that it has no floor apart from a narrow ledge adjoining the walls. The participant stands on a plank over the precipice and looks down to another room that is approximately 6 meters below and sees some furniture. The utility of this environment is that the expected responses are clear—people should show signs of anxiety. Because presence is operationally the extent to which people respond realistically, anxiety in this context is a sign of presence.

The pit environment was inspired by the famous "visual cliff" experiments of Eleanor Gibson and Richard Walk,[7] who investigated depth perception in different animals, in particular human babies. This was to examine whether they learn to avoid precipices through experience or whether it's an innate property of how the visual system interprets the patterns of light associated with depth—evidence, especially from animal trials, suggests the latter interpretation.

Mel Slater and his colleagues first used such an environment in virtual reality for an experiment that tested a method of locomotion based on walking-in-place.[8] Michael Meehan and his colleagues also used this experiment with physiological measures.[9] Most relevant to the work reported in this article was its use in a study of the impact of rendering quality.[6]

Rendering the Pit Room

In our new implementation, the room consisted of 1,535 polygons. The VE was displayed in stereo through a Virtual Research V8 head-tracked head-mounted display (HMD), which had 2 × 640 × 480 resolution. We used a Polhemus Fastrack tracking system for the head and handheld wand.

We rendered the VE using a parallel ray-tracing implementation run across a cluster of five dual-processor Xeon 3.2-GHz workstations; four were used as a rendering cluster and one as a master. We rendered the VE using two rendering methods both implemented through ray tracing (or ray casting). The first rendering method used an illumination model similar to OpenGL per-pixel local illumination without shadow effects (RC), and the second method used recursive ray tracing that included rendering shadows and reflections (RT). The render cluster performed ray-polygon intersections using a 4-ray SIMD (single instruction, multiple data) intersection method. The master workstation was responsible for control of the rendering cluster, the HMD, and the two trackers (head and right hand). The master used basic inverse kinematics to determine avatar pose and also issued render tasks. The render tasks were created by a simple tiling of the display surface across the HMD's two screens. Client workstations requested render tasks from the master using demand-driven scheduling. The cluster was configured to consistently deliver a stable frame rate (15 fps) that was kept fixed for the two rendering methods.

A separate workstation recorded electrodermal activity and ECG physiological data from a TTL ProComp Infiniti encoder during the experiments.

Experimental Design

We recruited 33 participants for the experiment via advertising around the University College London campus. They were split arbitrarily into two groups of RT (n = 17) and RC (n = 16). There were 15 females, 7 in the RC group and 8 in the RT group. Members of RT experienced the pit room rendered using RT and then the environment again

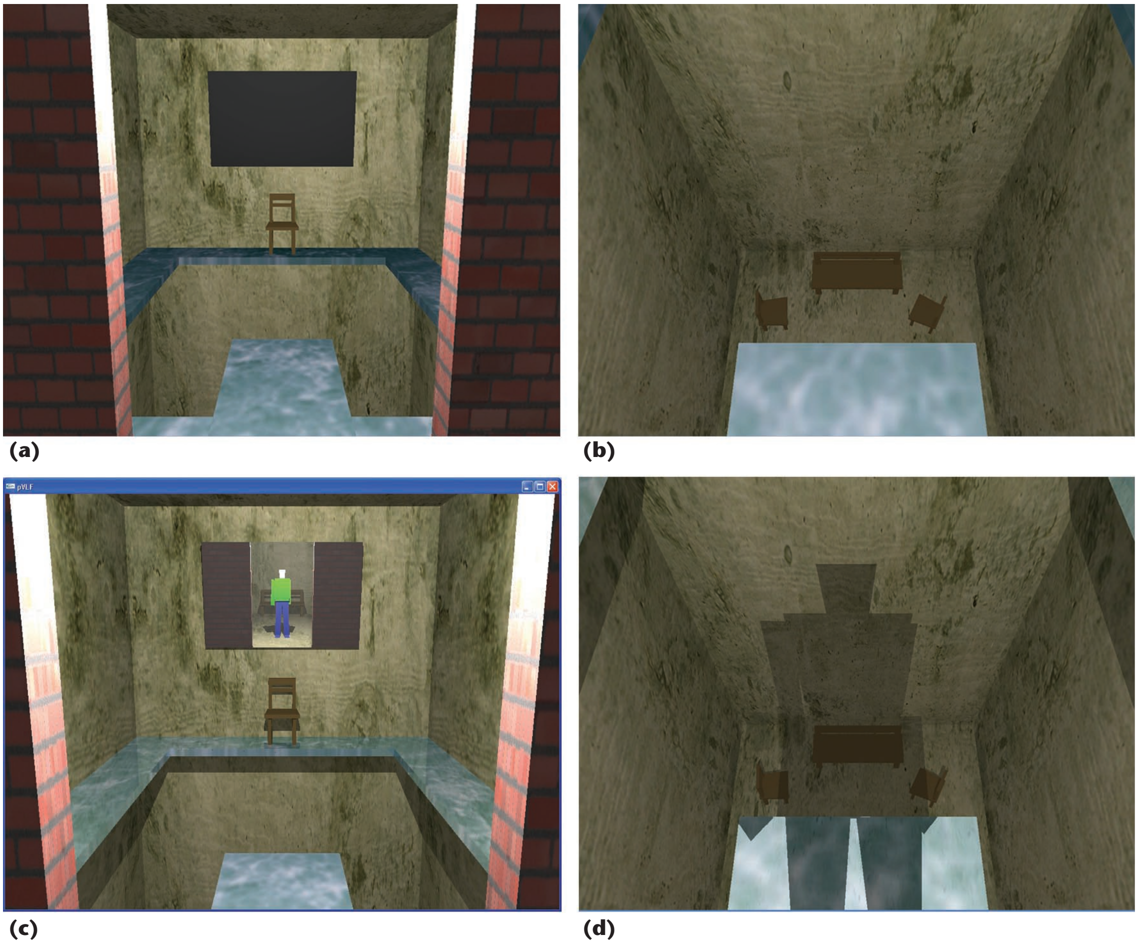

Figure 1. The pit room scene. (a) Looking into the pit room with ray casting. (b) Looking down into the pit with ray casting. (c) Looking into the pit room with real-time recursive ray tracing. (d) Looking down into the pit with real-time recursive ray tracing.

rendered with RC. Members of RC experienced the RC-rendered pit room and then the RT-rendered pit room. The experimental design was, therefore, both between-groups and within-groups. If we consider only the first exposure results, then it's between-groups. If we consider both exposures and compare between them, it's between-groups. We doubted the validity of a within-groups design. Because the first exposure has an influence on the results of the second exposure, we recruited large-enough sample sizes so that a between-groups interpretation could also be given. We consider only the between-groups results here.

Procedures

At the start of the experiment, we gave participants an information sheet that explained the experimental procedures, possible dangers of using virtual reality equipment (for example, dizziness), and what we expected of them. We gave them a disclaimer form to sign, informed them that they were free to withdraw from the experiment at any time without giving reasons, and informed them that in any case they would receive the equivalent of US$10 for their participation. We invited them to share basic information such as age, gender, frequency of computer game playing, and prior experience with virtual reality. We explained that they would enter the virtual environment twice and answer a questionnaire after each exposure.

We then fitted the participants with equipment for physiological recording. They donned the HMD and entered the pit room shown in Figure 1 standing in the doorway, first facing away from the room with the pit. We told them to look around, gave them time to get comfortable with the apparatus, and told them they could walk around a small radius of less than one meter. We then asked them to relax for two minutes and made physiological baseline recordings. We then invited them to turn around and look directly into the pit room for three minutes. After this, they took off the HMD and completed a questionnaire about presence. They then put on the HMD again; the physiological recordings were continued, and once again we invited them to look into the pit room, seeing it rendered with the other rendering method. After three minutes, they came out of the environment and answered the same questionnaire.

A simple avatar represented each participant from an egocentric viewpoint—for example, if they looked down they would see their virtual body, legs, and feet. More importantly, in the RT condition they would see reflections and shadows of their

Question	1	7
1. Please rate your sense of being in the pit room, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place. I had a sense of being there in the pit room ...	At no time	Almost all the time
2. To what extent were there times during the experience when the pit room was the reality for you? There were times during the experience when pit room was the reality for me ...	At no time	Almost all the time
3. When you think back about your experience, do you think of the pit room more as images that you saw or more as somewhere that you visited? The pit room seemed to be more like ...	Images that I saw	Somewhere I visited
4. During the experience, which was strongest on the whole, your sense of being in the pit room or of being in the real world of the laboratory? I had a stronger sense of …	Being in the pit room	Being in the lab
5. During the experience, did you often think to yourself that you were just in a laboratory, or did the pit room overwhelm you? During the experience I was thinking that I was really in the VR laboratory ...	Most of the time	Rarely
6. How much did you behave within the pit room as if the situation were real? I responded as if the situation were real …	Not at all	Very much
7. How often did you find yourself automatically behaving within the pit room as if it were a real place? I responded as if it were a real place …	Never	Almost all the time
8. How much did you deliberately behave within the pit room as if it were a real place? I deliberately responded as if it were a real place…	Never	Almost all the time
9. How much was your emotional response in the pit room the same as if it had been real? My emotional response in the pit room was the same as if it had been real …	Never	Almost all the time
10. How much were the thoughts you had within the pit room the same as if it had been a real situation? My thoughts within the pit room were the same as if it had been real …	Never	almost all the time
11. To what extent were your physical responses within the pit room (for example, heart rate, blushing, sweating) the same as if it had been a real situation? (In this case, if in such a real situation you would have had no or few such physical responses and also within the pit room you had no or few physical responses, then your answer should be closer to 7 than to 1). My physical responses within the pit room were the same as if it had been real …	Never	Almost all the time

Figure 2. Pit room questionnaire. Participants responded to each question on a 7-point Likert scale.

avatar, and these would move dynamically as the person moved. Participants held a tracked wand in their right hand. As they moved their real arm holding the wand, they would see, in reflections and shadows, their virtual arm move in response. This was the major difference between RC and RT. Figure 1c shows a reflection of the avatar in a mirror opposite the door to the pit room.

Questionnaire Results

The questionnaire we gave the participants immediately after their experience in the pit room consisted of 16 questions, of which 11 related to presence (see Figure 2). Each question was on a 7-point Likert scale, where we asked the participant to choose a number between 1 and 7 indicating the strength of their agreement with the statement in the question. For each question except question 4, a higher score means higher reported presence in the pit room. For purposes of analysis, we reversed the direction of question 4 so that all scores point in the same direction.

We found the mean score of each of the questions for the RC group (n = 17) and the RT group (n = 16) (see Table 1). This results in 11 pairs of values, one pair for each of the questions. If we look at each question, we find that the mean is higher for RT in 9 out of 11 questions. For question 3 the difference is significant using a nonparametric Kruskal-Wallis test ($P = 0.019$) and the results are similar for question 4 ($P = 0.007$).

Overall, the evidence suggests that the participants in RT reported a higher level of presence than those in RC. In particular, participants remember the pit room as a place they had visited (Q3), and their sense of being in the pit room was stronger in direct comparison to the real world of the laboratory (Q4). It's noteworthy that in earlier work looking at questions that best discriminated people's reported presence in virtual and real environments, Q3 was the best discriminator and Q4 the second best among those questions that had counterparts to the ones here.[10]

Physiological Recordings

Questionnaires provide insight into the participants' conscious thoughts and feelings, but they can't reveal what happened at a deeper level. For this purpose we employ physiological measures, in particular electrodermal activity (EDA) and ECGs (see the sidebar "Measuring Electrodermal and Cardiac Activity," next page).

We recorded EDA during the two-minute baseline and throughout the rest of the experiment, and derived the number and amplitude of skin conductance responses (SCR). We collected valid EDA data for 27 participants. In Table 2 we compare each person's baseline SCR rate (that is, the number of SCRs per 10 seconds) with their rate during the experimental condition for the RC and RT conditions. We made the same comparison for the SCR's mean amplitude. In each case we find that for RC there's no significant difference between the baseline and experimental condition, but for RT there's a significant difference, with experimental condition values on the average higher than baseline values. This indicates that between the baseline and experimental condition people's arousal and orienting responses were on the average greater in the case of RT but not in the case of RC.

We computed the significance levels by paired *t*-tests (comparing each person's baseline result with their experimental result; see Table 2). This test requires the set of differences for each person for each comparison to follow a normal distribution. Jarque-Bera tests for normality on each set don't reject the hypotheses of normality. The apparent difference between the ray casting mean baseline and the ray tracing mean baseline isn't significant (P = 0.79 using a nonparametric Wilcoxon rank sum test).

The EDA analysis shows a differential impact of RC and RT in terms of overall arousal, but it does not provide information about corresponding emotional significance. For that we turn to the ECG recordings, which were available for all 33 participants. First we consider heart rate and heart rate variability—in particular the heart rate (HR) divided by its standard deviation (*SDHR*). This quantity ($S = HR/SDHR$) increases with higher heart rate or lower heart rate variability. We found that *S* was significantly higher for the RT group than for the RC group, controlling for other factors, in particular for *S* during the baseline period and also for gender (see sidebar "Regression Analysis Variables," page 83).

A second ECG-derived parameter is the number of intervals of successive normal-to-normal intervals greater than 50 milliseconds (NN50) score. Lower values of NN50 would be a sign of higher mental stress. Indeed we find that NN50 is significantly lower for the RT group compared to the RC group in the experimental condition (taking into account the baseline NN50 scores) but not in the baseline condition (see sidebar "Regression Analysis Variables").

Finally, there is evidence of a difference between RC and RT with respect to the HF_{norm} parameter (a frequency domain measure). There is no significant difference between the groups for the baseline values (P = 0.17), but with respect to the

Table 1. Mean ± standard deviation of the questionnaire scores for the RC Ray Casting and Ray Tracing group.

Question	RC *n* = 17	RT *n* = 16
1	4.5 ± 1.6	4.6 ± 1.3
2	3.1 ± 1.6	3.2 ± 1.2
3*	2.9 ± 1.2	4.2 ± 1.5
4*	3.8 ± 1.6	5.3 ± 1.3
5	3.6 ± 2.0	3.5 ± 1.9
6	4.1 ± 2.0	3.9 ± 1.5
7	4.1 ± 1.7	4.4 ± 1.5
8	3.5 ± 1.6	3.7 ± 1.7
9	4.2 ± 1.5	4.8 ± 1.3
10	4.2 ± 1.8	4.5 ± 1.5
11	3.6 ± 1.6	3.8 ± 1.6

**Indicates significant difference*

Table 2. Rate and mean amplitude of skin conductance responses (SCRs).

	Mean number of SCRs per 10 seconds	Mean amplitude of SCRs (μS)
Ray casting (*n* = 14)		
Baseline	1.16 ± 0.87	0.25 ± 0.11
Experiment	1.35 ± 1.10	0.30 ± 0.19
Significance Level	0.36	0.13
Ray tracing (*n* = 13)		
Baseline	1.00 ± 0.65	0.24 ± 0.12
Experiment	1.38 ± 0.84	0.31 ± 0.17
Significance Level	0.03	0.01

Table 3. Linear regression for S_{exp} = HR/SDHR. Multiple correlation $R^2 = 0.72$, $F = 24.51$ on (3, 29) d.f., $P = 4.2\times10^{-8}$. A Jarque-Bera test does not reject the hypothesis that the residual errors of the fit follow a normal distribution (P = 0.22).

Parameter	Estimate	P
Constant	2.99	—
S_{base}	0.75	0.0000
Gender (F = 1, M = 0)	–0.24	0.0203
Condition (RC = 0, RT = 1)	2.70	0.0097

Measuring Electrodermal and Cardiac Activity

The physiological measures recorded during this experiment were electrodermal activity (EDA)[1] and electrocardiogram (ECG). We fitted participants with a ProComp Infiniti (Thought Technology) physiological recording device that recorded the ECG (256 Hz) and skin conductance (32 Hz). We placed electrodes on the palmar areas of the index and middle fingers of the left hand to record electrodermal activity. We also placed electrodes on the left and right collar bones and the lowest left rib to record ECG.

EDA measures changes in arousal through changes in skin conductance caused by sweat levels. Figure A shows an example recording during the baseline period for an arbitrarily chosen participant. An important derived measure of interest is the number of skin conductance responses (SCR), which reflect transient sympathetic arousal, either spontaneous or in response to events, specifically responses to changes in the environment and events or surprises. Examples of SCRs are shown in Figure A. We defined SCRs to be local maxima that had amplitude of at least 0.1 μS, in a period not exceeding five seconds from the start of the SCR to its maximal point. (There is no standard definition, and we used these criteria as a compromise among several in the literature.) The amplitude refers to the maximum level reached compared to the start of the SCR. Of interest are both the number and amplitude of SCRs, and we also refer to the SCR rate as the number of SCRs per 10 seconds. Such SCRs were identified in an offline program written in Matlab.

Figure B shows a sample ECG series. From the raw ECG, the so-called QRS complexes are computed offline. These QRS complexes determine the time between heart contractions—the RR intervals. An NN interval refers to normal-to-normal intervals, where nonnormal beats such as extra systoles are not taken into account. From the ECG signal and QRS complexes, a number of parameters can be derived. In our analysis we used the following time domain measures:[2]

- HR—heart rate in beats per minute (bpm)
- SDHR—heart rate variability as measured by the standard deviation of heart rate (bpm)
- NN50—number of intervals of successive NN intervals greater than 50 milliseconds.

Generally episodic higher HR and lower SDRH indicate either exercise or mental stress. NN50 is also another indicator of heart rate variability—lower values indicate lower variability.

In the frequency domain, we examined low-frequency components (LF, 0.1 Hz) and high-frequency components (HF, 0.15–0.4 Hz). These indicate mental stress when the LF component increases and the HF component decreases.

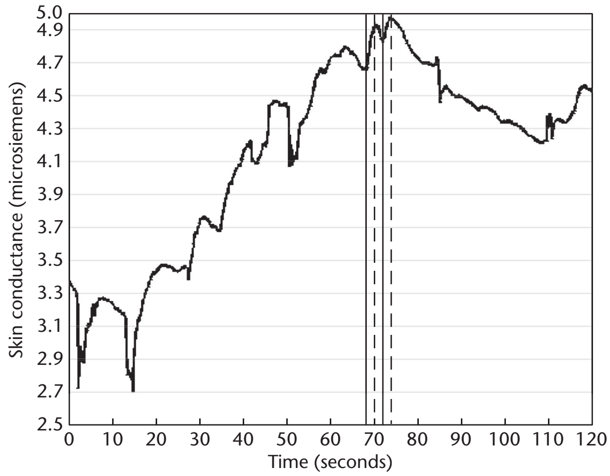

Figure A. Skin conductance during the baseline period for one participant. The two solid vertical lines indicate two of the detected skin conductance responses, and the dashed lines their maxima.

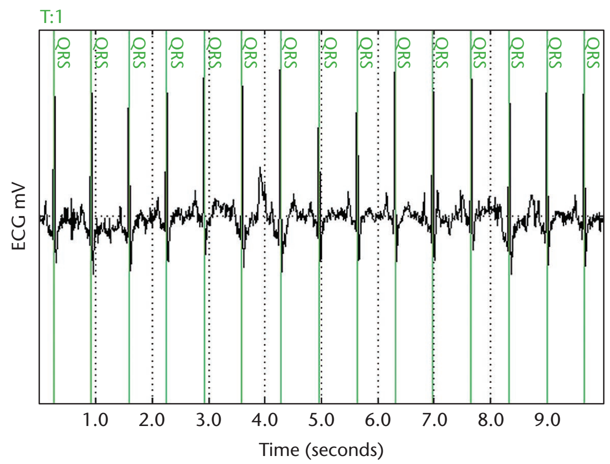

Figure B. First 10 seconds of electrocardiogram wave form during the baseline for one participant. The QRS complexes determine the time between heart contractions and are computed offline.

Moreover, during dynamic exercise the heart rate changes but the HF component doesn't change significantly; hence, a change in HF together with changes in HR and SDHR indicates mental stress.

References

1. W. Boucsein, *Electrodermal Activity*, Plenum Press, 1992.
2. C. Guger et al., "Heart-Rate Variability and Event-Related ECG in Virtual Environments," *Presence 2004: 7th Int'l Conf. Presence*, Universidad Politecnica de Valencia, 2004, pp. 240–245.

experimental condition HF_{norm} is lower for RT than for RC (P = 0.07). This provides further evidence that changes in heart rate and heart rate variability were due to mental stress rather than simply physical exercise. In any case the participants' tasks in both RC and RT were the same, so it's unlikely that any observed differences would be due to differences in physical effort.

The substantial difference between the RC and RT conditions was that RT had shadows and reflections (especially shadows of the virtual body) and RC did not. Other than that, they both used the Phong lighting model and were texture mapped. The results of our experiment suggest that this difference led to different levels of anxiety and different levels of reported presence between the two groups. The analysis of physiological recordings suggests that participants became more aroused relative to their own baseline during the RT condition (EDA analysis) and that the RT group as a whole became more stressed (allowing for differences in baseline) than the RC group (ECG analysis).

We obtained further insight on this from the free-form interviews after the conclusion of the experiment, when participants reflected on their experiences of both conditions. Of the 24 who expressed an opinion on the impact of shadows and reflections on their sense of presence, 17 expressed positive opinions. Of the seven who said the shadows and reflections had a negative impact, the common view was that it was because the reflection didn't look like themselves. For example, one comment was,

> The reflection was there too, but obviously looking at a mirror you expect to see yourself, not a blue Lego. So I think that it sort of spoiled it." Another participant who thought positively of the impact said: "I rather like the shadow which kind of loomed into the space below me when I looked down.... I reacted quite strongly to the figure, my reflection. Although, once I realized what it was, I mean, initially my response was mild fear as to what it was, and I couldn't you know, and I couldn't identify it as being a reflection because it didn't look like that. But once I figured it out ... it did enhance the realism of the environment in terms of the reflection, and looking down and being able to see your own reflection, kind of seeping into space, your shadow. From that point of view it heightened the realism of the space, even though it was abstract in itself and quite blocky.

The take-home message of this experiment is that improved visual realism might enhance realistic

Regression Analysis Variables

We carried out a regression analysis for the response variable S_{exp}, which is S during the experimental period. To eliminate the effect of differences between individuals we use S_{base}, which is S during the baseline period, as an explanatory variable, and the experimental condition (C) treated as a binary variable (ray casting—RC—as 0, ray tracing—RT—as 1) as the independent variable. In addition, we found another explanatory variable, gender, to be significant. This three-variable model led to a highly significant fit with correlation R^2 = 0.72, which means that 72 percent of the variation in S_{exp} could be explained by the variation in S_{base}, C, and gender. In particular, S_{exp} varies positively with S_{base}, is lower for females, and is higher for the RT condition. It's important to note that there's no significant difference in S_{base} between the RT and RC conditions, so the difference in S_{exp} can only be due to the impact of different rendering styles.

NN50 is a count variable that should be modeled by a Poisson distribution as events (adjacent NN50 intervals differing by more than 50 milliseconds) occurring randomly in time. The appropriate regression model to use is Poisson log-linear regression. Considering the baseline measurements only, there's no difference in NN50 between the RC and RT; however, there is for the experimental period, as shown in Table A.

HF_{norm} is a frequency domain measure, and we tested the difference between RC and RT. When comparing the baseline measures there was no significant difference between the groups (P = 0.14). For the experimental period the RT group had a lower score than the RC group, just outside the conventional 5 percent limit (P = 0.066), in each case using one-way analysis of variance; the hypothesis of normality of the residual errors was not rejected in either case by a Jarque-Bera test. The best regression model fit is obtained using the independent factor C, the explanatory variables NN50 for the baseline, and age. This model has R^2 = 0.41, and HF_{norm} has a negative coefficient for RT with significance level P = 0.06.

References

1. N. Breslow, "Extra-Poisson Variation in Log-Linear Models," *Applied Statistics*, vol. 33, no. 1, 1984, pp. 38–44.

Table A. Poisson log-linear regression for NN50. Deviance for the overall model = 33.1 on 30 d.f. (*P* = 0.32) based on the method of fit described by Norman Breslow.[1]

Parameter	Estimate	*P*
Constant	–2.5545	
Baseline NN50/baseline time	4.4986	0.0000
Condition (RC = 0, RT = 1)	–0.5825	0.0097

behavioral response. Of course, the results here apply to a virtual environment designed to result in anxiety, and we don't know whether the results extend to more mundane applications. Moreover, although it's likely that the dynamic shadows and reflections caused the changed responses, we can't disambiguate these from the general improvement in visual quality that results from recursive ray tracing—though given the results from another experiment[8] it's unlikely that visual quality alone can account for the differences found here.

In our current experimental work we separate the effects of dynamic shadows and reflections from the effect of improved visual realism by employing a full real-time global-illumination solution rather than only recursive ray tracing, and in a nonstressful environment. In addition, overcoming one problem discussed previously, we also use highly realistic avatars rather than the simple blocky avatar used here.

Acknowledgments

This research was carried out under the UK Engineering and Physical Sciences Research Council (EPSRC) Project Presence in a Virtual Light Field (EP/C511824/1). Mel Slater's contribution was also supported through the EU Future and Emerging Technologies (FET) Integrated Project Presenccia Contract Number 27731. We thank Dr. Christoph Guger (Presenccia project) for his advice on the analysis of the ECG data.

References

1. M.V. Sanchez-Vives and M. Slater, "From Presence to Consciousness through Virtual Reality," *Nature Reviews Neuroscience*, vol. 6, no. 4, 2005, pp. 332–339.
2. M. Mori, "Bukimi No Tani (The Uncanny Valley)," *Energy*, vol. 7, no. 4, 1970, pp. 33–35 (in Japanese).
3. C. Hendrix and W. Barfield, "Presence within Virtual Environments as a Function of Visual Display Parameters," *Presence-Teleoperators and Virtual Environments*, vol. 5, no. 3, 1996, pp. 274–289.
4. M. Slater, M. Usoh, and Y. Chrysanthou, "The Influence of Dynamic Shadows on Presence in Immersive Virtual Environments," *Selected Papers Eurographics Workshops Virtual Environments 1995*, Springer, 1995, pp. 8–21.
5. K. Mania and A. Robinson, "The Effect of Quality of Rendering on User Lighting Impressions and Presence in Virtual Environments," *ACM Siggraph Int'l Conf. Virtual Reality Continuum and its Applications in Industry*, ACM Press, 2004, pp. 200–205.
6. P. Zimmons and A. Panter, "The Influence of Rendering Quality on Presence and Task Performance in a Virtual Environment in Virtual Reality," *Proc. IEEE Virtual Reality Conf.* (VR 03), IEEE CS Press, 2003, pp. 293–294.
7. E.J. Gibson and R.D. Walk, "The Visual Cliff," *Scientific Am.*, vol. 202, no. 4, 1960, pp. 64–72.
8. M. Slater, M. Usoh, and A. Steed, "Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality," *ACM Trans. Computer-Human Interaction*, vol. 2, no. 3, 1995, pp. 201–219.
9. M. Meehan et al., "Physiological Measures of Presence in Stressful Virtual Environments," *ACM Trans. Graphics*, vol. 21, no. 3, 2002, pp. 645–652.
10. M. Usoh et al., "Using Presence Questionnaires in Reality," *Presence-Teleoperators and Virtual Environments*, vol. 9, no. 5, 2000, pp. 497–503.
11. N. Breslow, "Extra-Poisson Variation in Log-Linear Models," *Applied Statistics*, vol. 33, no. 1, 1984, pp. 38–44.

Mel Slater *is a research professor at the Catalan Institute for Research and Advanced Studies (ICREA) at the University of Barcelona, and a professor of virtual environments at University College London. His research interests include global illumination algorithms for virtual reality and scientific understanding of how people respond to virtual reality. Slater has a DSc from the University of London. Contact him at mel.slater@ub.edu.*

Pankaj Khanna *currently is a consultant in financial services and maintains links to academia as an honorary research fellow at University College London. His research interests include real-time physically based rendering and virtual reality. Contact him at p.khanna@cs.ucl.ac.uk.*

Jesper Mortensen *is a graphics programmer at Geomerics working on real-time global illumination for the games industry. His research interests include rendering algorithms and global illumination. Mortensen has an MSc in vision, imaging, and virtual environments from University College London. Contact him at jesper@geomerics.com.*

Insu Yu *is a finishing PhD student at University College London. His research interests include real-time global illumination on the GPU, especially based on spherical harmonics. Yu has an MSc in vision, imaging, and virtual environments from University College London. Contact him at i.yu@cs.ucl.ac.uk.*