Visual Realism Enhances Realistic Response in an Immersive Virtual Environment

Mel Slater[1,2] , Pankaj Khanna[1], Jesper Mortensen[1], Insu Yu[1]

[1]Department of Computer Science, University College London, UK
[2]ICREA @ Universitat Politècnica de Catalunya, Centre de Realitat Virtual, Departament de LSI, Barcelona, Spain

Abstract

Does greater visual realism induce greater presence of participants in immersive virtual environments? Presence refers to how realistically participants respond to the environment as well as their subjective sense of being in the place depicted by the VE. 33 people were exposed for 3 minutes to a virtual environment depicting a precipice in a head-tracked head-mounted display system. 17 of them saw the environment rendered with real-time recursive ray tracing (RT) that included shadows and reflections of their virtual body and the remainder experienced the same environment rendered with ray casting (RC) which did not include shadows and reflections. Participants completed a presence questionnaire immediately after their experience, and physiological responses (skin conductance and electrocardiogram) were recorded throughout. The results show that subjective presence was higher for the RT environment than for the RC one and that higher stress was induced in the RT environment compared to the RC one.

Keywords

Virtual environments, presence, visual realism, real-time ray tracing, shadows, reflections, virtual body, avatar

Introduction

Participants in an immersive virtual environment interact with the scene from an egocentric point of view, where their bodies appear to be located, rather than from the outside as if through a window. People interact through normal body movements, such as head-turning, reaching, bending, and within the limitations of the tracking available, are able to move through the environment, or effect changes within it, in more or less natural ways. A major application of such systems is for rehearsal or training for situations that would be too dangerous or impractical to carry out in real life, such as vehicle simulators. Another major application is in the realm of psychotherapy where patients may experience anxiety-provoking events within the safe knowledge that nothing real is happening, but where nevertheless their normal anxiety responses are induced and ultimately reduced through repeated exposure. What is common to all of these is that success of the application relies on the participants responding realistically to events and situations within the virtual environment as if these were real – for if not, any learning achieved would not be transferable 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, behavioural and physiological (such as changes in heart rate).

Here we consider the impact of the level of *visual realism* on presence. Visual realism has two components – geometric realism (the virtual object looks like the real object that it represents) and illumination realism (referring to the fidelity of the lighting model). Both types have static and dynamic aspects. An object might statically look realistic, but its dynamic changes through time may not – for example, an object representing a human may look but not behave realistically. Similarly good static lighting may be achieved with a method such as radiosity which pre-computes all diffuse interreflections in the environment, but shadows may not move when the corresponding object is moved. Here we focus in particular on the impact of illumination, such as can be achieved by real-time ray tracing.

One hypothesis says that visual realism may not be important for presence. This is based on the idea that the human perceptual system works in a top down manner, building apparently complete representations of an environment from a few minimal cues ([1] and references therein). 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 head-tracked system. Indeed it has been argued that should the level of displayed realism improve 'too much' then this could lead to degradation in response, since human observers will be more likely to notice small imperfections. This is Masahiro Mori's 1970 'Uncanny Valley' hypothesis - that improvements in quality might result in improvements in response up to a point after which there may be a sudden dip in response due to the problem of defect magnification.

However, the UV hypothesis is very convenient where, at least in the context of real-time rendering for immersive VEs, it has not until recently been feasible to produce high quality visual rendered environments – real time global illumination methods have not been possible. Moreover, the evidence such as there is, has not been clear regarding the impact of visual quality on presence. Here we show that when recursive real-time ray tracing is used to render an immersive VE, appropriate anxiety amongst participants is significantly higher 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 that has been used several times before in tests for presence.

Background

Early papers that addressed the issue of the impact of visual realism on presence, for example [2], compared presence amongst environments with different levels of visual detail and different levels of pictorial realism. Greater reported presence was found for higher visual realism in both senses. By 'reported presence' we mean the sense that participants had of being in the place depicted in the virtual environment, as assessed by questionnaires. Another study compared reported and behavioral presence across rendering styles that either did not support shadows, supported static shadows, or supported dynamically changing shadows. Greater behavioural and reported presence was found for static shadows compared to no shadows, and dynamic shadows compared to static shadows [3]. In that case dynamic shadows were computed using an adaptation of the Binary Space Shadow Volume Binary Space Partition algorithm.

In a study that used two levels of radiosity and also flat shading, no difference in reported presence using a presence questionnaire was found between the three conditions [4]. Finally, in an experiment that was set in the 'pit room' environment, similar to the one that we used in our experiment, the scene was displayed at various levels of illumination realism (wire frame, without and with textures, and with radiosity) [5]. In that experiment physiological measures were recorded as well as questionnaire responses. It was found that all subjects exhibited significantly increased heart rate when they encountered the pit, although there were no significant differences in heart rate or reported presence between the different rendering conditions.

Results to date therefore show no clear message about the impact of the level of visual realism on reported or behavioural presence. There is some evidence to support both conclusions that type of rendering may or may not impact presence. Moreover we believe that many previous studies were flawed, based on within-group designs (see Inset 1), where participants would experience all experimental conditions (the different levels of visual realism) and therefore would obviously realise the purpose of the experiment. It is important to note that this does not apply to [5] which is the study that is most directly comparable with the one reported here.

Our study differs in three important ways from those reported above. First, we use real-time recursive ray tracing, which although not full global illumination correctly simulates light transport between specular surfaces, and therefore generates dynamically changing reflections and shadows for point light sources. In particular it generates reflections and shadows of the virtual body that represents the participant immersed in the (head-mounted display generated) virtual reality. Although [3] had dynamic shadows, these were limited to a single object on a flying trajectory, and did not include shadows of the virtual body of the participant. 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 provide an analysis of the physiological responses of the participants, in particular we carry out electrocardiogram (ECG) analysis in order to determine levels of stress. Third, our study is between-groups, which means that the results are based on each participant's experience of only one condition (recursive ray tracing or ray casting) and therefore cannot be biased by understanding the purposes of the study.

The Pit Room Environment

The scene used in our experiment was a variation of the so called 'pit room'. In such pit room scenarios the participant first enters into a virtual training room. This is typically an ordinary room with some furniture where the participant accustoms to the environment, and maybe learns to carry out some tasks, depending on the requirements of the particular experiment. Then he or she is required to move into an adjoining room through an open door, and this room at first sight seems normal, but then is seen to have no floor apart from a narrow ledge adjoining the walls. The participant stands at first on a plank over the precipice and can look down to another room which is approximately 6m below, and can see some furniture. The utility of this environment is that the expected responses are clear: people should show signs of anxiety. Since presence operationally is the extent to which people respond realistically, then anxiety in this context is a sign of presence.

The pit environment was inspired by the famous 'visual cliff' experiments of Gibson and Walk [6] who were investigating depth perception in different animals, in particular human babies. This was in order to examine whether they learn to avoid precipices through experience or whether this is an innate property of how the visual system interprets the particular patterns of light associated with depth - the evidence, especially from animal trials, suggesting the latter interpretation.

Such an environment was first used in virtual reality in [7] in an experiment that tested a method of locomotion based on walking-in-place. It was also used with physiological measures in [8]. Most relevant to the work reported in this paper was its use in [5] in a study of the impact of rendering quality.

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 has 2× 640×480 resolution. The tracking system used for the head and hand-held wand was a Polhemus FASTRACK.

The VE was rendered using a parallel ray-tracing implementation run across a cluster of five dual-processor Xeon 3.2Ghz workstations . Two rendering methods were used both implemented through ray tracing (or ray casting). The first, an illumination model similar to OpenGL per-pixel local illumination without shadow effects (RC), and a recursive ray-tracing method capable of rendering shadows and reflections (RT). Ray-polygon intersections were performed using a 4-ray SIMD intersection method. Control of the rendering cluster, HMD and the two trackers (head and right hand) were performed by a master workstation that used basic inverse-kinematics to determine avatar pose and issue render tasks. These render tasks were created by a simple tiling of the display surface across the HMD's two screens. Render tasks were requested from the master by client workstations using demand-driven scheduling. The cluster was configured to be able to consistently deliver a stable frame-rate (15fps) that was kept fixed for the two rendering methods.

A separate workstation recorded electrodermal activity and ECG physiological data from a TTL ProComp Infiniti[1] encoder during the experiments.

Experimental Design

33 participants were recruited for the experiment by advertisement around the UCL 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[2]. Members of RT experienced first of all the pit room rendered using RT and then experienced the environment again but this time rendered with RC. Members of RC experienced first of all the RC rendered pit room and then the RT rendered pit room. The experimental design was, therefore, both between-groups and within-groups (Inset 1). If we consider only the first exposure results then it is between-groups. If we consider both exposures and compare between them it is within-groups. We considered it doubtful whether a within-groups design is valid: since the first exposure is bound to have an influence on the results of the second exposure. We therefore recruited large enough sample sizes so that a between-groups interpretation could also be given, and only the between-groups results are considered here.

Inset 1

A between-group design is one 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, etc) and this information can be used to further statistically equalise the groups in the later analysis. A between-group design has the advantage that the participants do not learn the purposes of the experiment, since they only experience one condition. It has the disadvantage that typically more people are needed than for a within-groups design for

[1] http://www.thoughttechnology.com/proinf.htm

[2] Our goal was 15 participants per group, but normally we over-recruit in order to avoid the problem of invalid physiological data.

good statistical results. A within-group design is one where each participant experiences every condition, the order of presentation randomised across the participants. This has the advantage that each person is compared to him- or herself, and also fewer are needed than for a between-groups design to achieve good statistical results.

In general the problem with within-groups designs in virtual reality experiments is that the conditions are not symmetric. In other words it is certain that the experience of one condition (e.g., 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 realise, or may think they realise, the purpose of the experiment, since they see all conditions. This could also bias their responses, especially if unwittingly they felt that they should give the kinds of answers that they believe would be pleasing to the experimenters. Finally, there is a question of adaptation. Having experienced the precipice once, their response may be different on the second exposure. In our earlier brief report of some aspects of the results of this experiment [9] indeed we showed that the responses were not symmetric comparing the between-group conditions only and the within-group conditions with regard to subjectively reported presence. In this paper we report only on the between-group condition from which the most reliable results can be obtained.

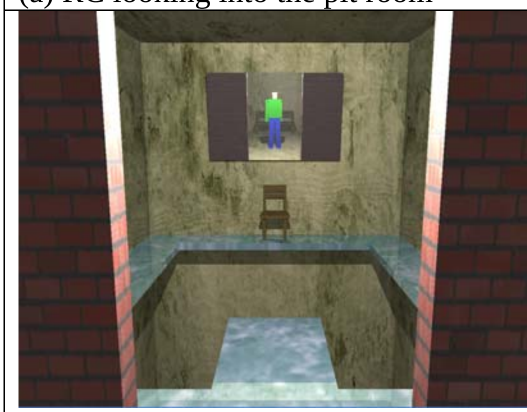
Procedures

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

Participants were then fitted with equipment for physiological recording, donned the HMD and entered the pit room shown in Figure 1, standing in the doorway, but first of all facing away from the room with the pit. They were told that they should look around, and time was given for them to become comfortable with the apparatus, and they could walk around within a small radius less than 1m. Then they were told to relax for two minutes and physiological baseline recordings were made. They were invited to turn around and would then be looking directly into the pit room. They were told to look around the room for a period of 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 invited to look into the pit room. The second time, however, they would see the pit room rendered with the other rendering method. After three minutes again they came out of the environment and answered the same questionnaire.

(a) RC looking into the pit room

(b) RC looking down into the pit



(c) RT looking into the pit room

(d) RT looking down into the pit

Figure 1 –The pit room scene (a-b) showing the RC version and (c-d) showing RT with shadows and reflections of the avatar.

Each participant was represented by a simple avatar that they could see 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 avatar, and of course these would move dynamically as the person moved. Each participant held a tracked Wand in their right hand. As they moved their real arm holding the tracker they would see, in reflections and shadows, their virtual arm move in response. This was the major difference between RC and RT – for example, Figure 1(d) shows a reflection of the avatar in a mirror opposite the door to the pit room.

Questionnaire Results

The questionnaire that was given immediately after their experience in the pit room consisted of 16 questions of which 11 related to presence (Inset 2). Each question was on a 7 point Likert scale, where the participant was asked to tick a number between 1 and 7 indicating the strength of their agreement with the statement in the question. The labels for the scores of 1 and 7 are shown in the questions in the sidebar. For each question except Q4 a higher score means higher reported presence in the pit room. For the purposes of analysis we reverse the direction of question 4 so that all scores point in the same direction.

Inset 2		
Question	**1**	**7**
1. Please rate your sense of being 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 time of 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 time of 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 thoughts with in the pit room were 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? In spite of my knowledge that the situation wasn't real I found myself behaving as if it were real…	never	almost all the time
11. To what extent were your physical responses within the pit room (e.g., heart rate, blushing, sweating, etc.) 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

We find the mean score of each of the questions for the RC group (n=17) and the RT group (n=16) (Table 1). This results in 11 pairs of values, one pair for each of the questions. If we look at each individual question we find that the mean is higher for RT in 9 out of the 11 questions. For question 3 the difference is significant using a non-parametric Kruskal-Wallis test (P=0.019), and similarly 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 the pit room tended to be remembered as a place that had been visited (Q3) and the sense of being in the pit room was stronger in direct comparison to the real world of the laboratory (Q4). It is noteworthy that in earlier work that looked at questions that best discriminated people's reported presence in virtual and real environments, that Q3 was the best discriminator, and Q4 the second best [10] amongst those questions that had counterparts to the ones here.

Table 1

Mean $\pm$ *Standard Deviation* of the questionnaire scores for the RC and RT group.
* indicates significant difference.

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

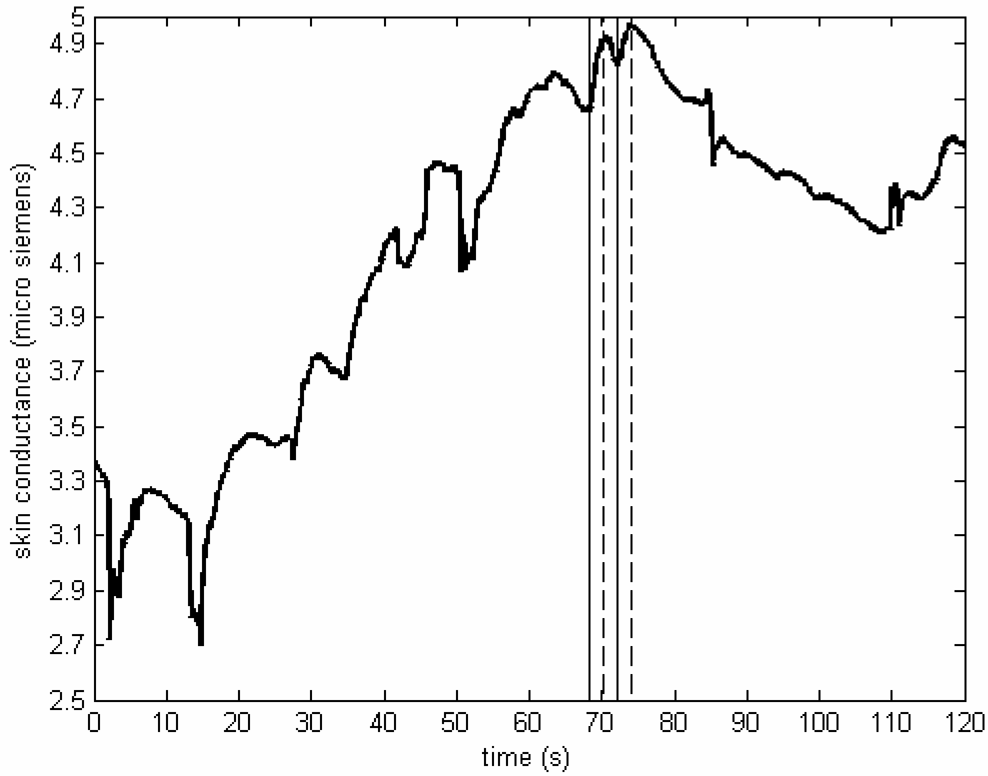


Figure 2 – Skin conductance during the baseline period for one participant. The two solid vertical lines indicate two of the detected Skin Conductance Responses, and the dashed lines their maxima.

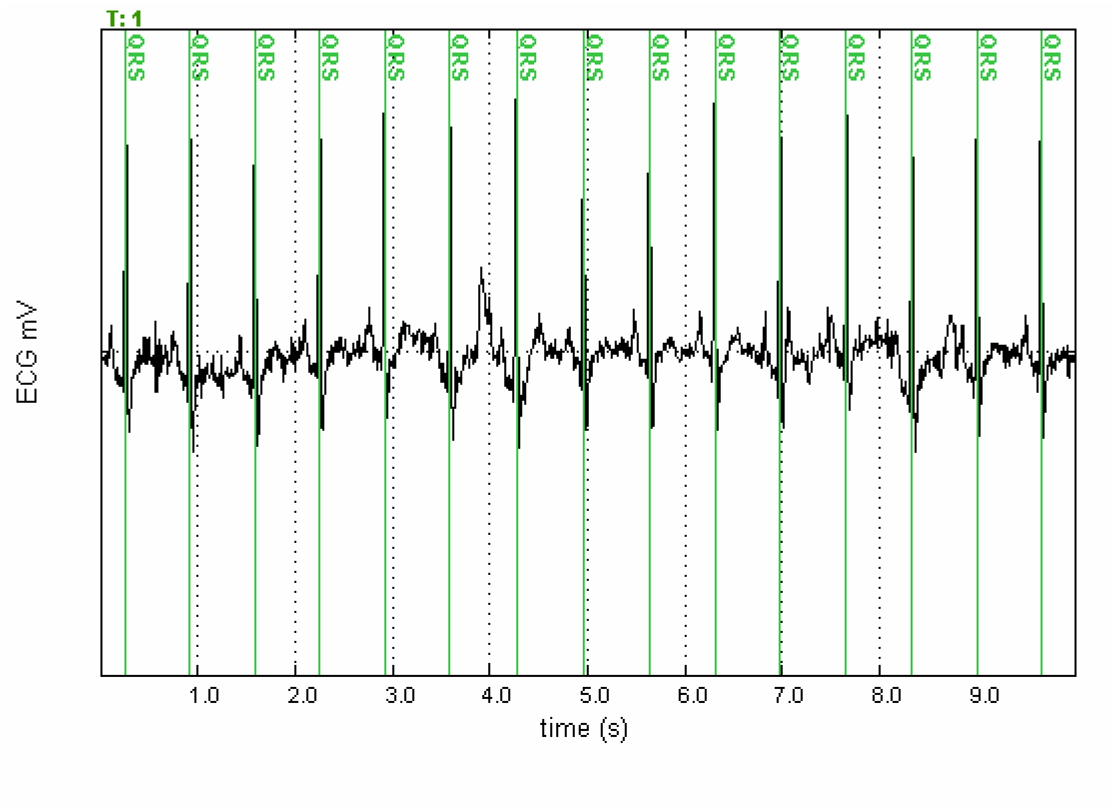


Figure 3 – First 10s of ECG wave form during the baseline for one participant. The QRS complexes are computed offline.

Inset 3
The physiological measures recorded during this experiment were electrodermal activity (EDA) [11] and electrocardiogram (ECG). The participants were fitted with a ProComp Infiniti (Thought Technology) physiological recording device that recorded the ECG (256Hz) and skin conductance (32Hz). Electrodes were placed on the palmar areas of the index and middle fingers of the left hand in order to record electrodermal activity. Electrodes were placed on the left and right collar bones and the lowest left rib in order to record ECG.

EDA measures changes in arousal through changes in skin conductance caused by sweat levels. An example recording during the baseline period for one arbitrarily chosen participant is shown in Figure 2. 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 the orienting response, that is responses to changes in the environment and events or surprises. Examples of SCRs are shown in Figure 2. Skin conductance responses (SCR) were defined to be local maxima that had amplitude of at least 0.1 μS and in a period not exceeding 5s from the start of the SCR to its maximal point. (There is no standard definition, and we used these criteria as being a compromise amongst 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 10s. Such SCRs were identified in an offline program written in MATLAB.

A sample ECG series is shown in Figure 3. From the raw ECG the so-called QRS complexes are computed offline[3]. These QRS complexes determine the time between heart contractions – the RR intervals. An *NN interval* refers to normal-to-normal intervals, where non-normal 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 have used the following time domain measures [12]:

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 ms.

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 - Low Frequency components (LF, 0.1 Hz) and High Frequency components (HF, 0.15-0.4 Hz) are examined. These indicate mental stress when the LF component increases and the HF component decreases. Moreover, during dynamic exercise the heart rate changes but the HF component does not change significantly, hence a change in HF together with changes in HR and SDHR indicates mental stress.

Physiological Recordings

Questionnaires provide some insight into conscious thoughts and feelings of the participants, but they cannot go ‘behind the scenes’ and reveal what was happening at a deeper level. For this purpose we employ physiological measures, in particular electrodermal activity EDA and electrocardiograms (ECG) (Inset 3).

Electrodermal activity (EDA) was recorded during the 2 minute baseline and throughout the rest of the experiment, and the number and amplitude of skin conductance responses (SCR) were derived. Valid EDA data was available for 27 participants. In Table 2 we compare each person’s baseline SCR rate (that is the number of SCRs per 10s) with their rate during the experimental condition, for the RC and RT conditions. The same comparison was made for the mean amplitude of the SCRs. In each case we find that for RC that there is no significant difference between the baseline and experimental condition, but for RT there is a significant difference, with the experimental condition values on the average higher than the 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.

[3] ECG analysis is carried out with gbsanalyze MATLAB package (Guger Technologies).

Table 2
Rate and Mean Amplitude of Skin Conductance Responses

The significance levels are computed by paired t-tests (comparing each person's baseline result with their experimental result). 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 do not reject the hypotheses of normality. The apparent difference between the ray casting mean baseline and the ray tracing mean baseline is not significant (P=0.79 using a non-parametric Wilcoxon rank sum test).

	Mean Number of SCRs per 10s	**Mean Amplitude of SCRs (μS)**
Ray Casting (n=14)		
Baseline: Experiment: Significance Level P:	1.16 ± *0.87* 1.35 ± *1.10* 0.36	0.25 ± *0.11* 0.30 ± *0.19* 0.13
Ray Tracing (n=13)		
Baseline: Experiment Significance Level P:	1.00 ± *0.65* 1.38 ± *0.84* 0.03	0.24 ± *0.12* 0.31 ± *0.17* 0.01

The EDA analysis show that there was a differential impact of RC and RT in terms of overall arousal, but it does not provide information about the 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 (Inset 4).

Inset 4
We carried out a regression analysis for the response variable Sexp, which is S during the experimental period. To eliminate the effect of differences between the individuals we use Sbase, which is S during the baseline period, as an explanatory variable, and the experimental condition (C) treated as a binary variable (RC as 0, RT as 1) as the independent variable. In addition another explanatory variable, gender, was found to be significant. This three variable model led to a highly significant fit with correlation R^2 = 0.72, which means that 72% of the variation in Sexp could be explained by the variation in Sbase, C, and gender. In particular, of course Sexp varies positively with Sbase, is lower for females, and higher for the RT condition. It is important to note that there is no significant difference in Sbase between the RT and RC conditions, so that the difference in Sexp can only be due to the impact of the 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 ms) occurring randomly in time. The appropriate regression model to use is Poisson log-linear regression. Considering the baseline measurements only there is no difference in NN50 between the RC and RT, however, there is for the experimental period, as shown in Table 4.

HFnorm is a frequency domain measure and the difference between RC and RT was tested. When comparing the baseline measures there was no significant difference between the groups (P = 0.14). For the experimental period the RT group had lower score than the RC group, just outside the conventional 5% limit (P = 0.066), in each case using one way ANOVA (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 in addition to the independent factor C, the explanatory variables NN50 for the baseline, and age. This model has R^2 = 0.41 and HFnorm has a negative coefficient for RT with significance level P=0.06.

Table 3

Linear Regression for Sexp = HR/SDHR

Multiple Correlation $R^2 = 0.72$, F = 24.51 on (3,29) d.f., P = 4.2×10^{-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	
Sbase	0.75	0.0000
Gender (F=1,M=0)	-2.40	0.0203
Condition (RC=0, RT=1)	2.70	0.0097

A second ECG derived parameter is the 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 (Inset 4).

Table 4

Poisson Log-linear Regression for NN50

Deviance for the overall model = 33.1 on 30 d.f. (P = 0.32)

The method of fit used is described in

Breslow, N., Extra-Poisson Variation in Log-Linear Models.

Applied Statistics, 1984. 33(1): p. 38-44.

Parameter	Estimate	P
Constant	-2.5545	
Baseline NN50/baseline time	4.4986	0.0000
Condition (RC=0, RT=1)	-0.5825	0.0073

Finally there is some evidence of a difference between RC and RT with respect to the HFnorm parameter. There is no significant difference between the groups for the baseline values (P=0.17), but with respect to the experimental condition the HFnorm is lower for the RT than for the RC (P = 0.07). This provides further evidence that the changes in heart rate and heart rate variability were due to mental stress rather than simply to physical exercise. In any case the tasks of the participants in both RC and RT were the same, so it is unlikely that any observed differences would be due to differences in physical effort.

Conclusions

The substantial difference between the RC and RT conditions was the fact that one had shadows and reflections (especially of course shadows of the virtual body) and the other 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 between the two groups and different levels of reported presence. The analysis of the 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).

Some further insight on this was obtained from the free-form interviews after the conclusion of the experiment, when participants had the chance to reflect on their experiences of both conditions. Of the 24 who expressed an opinion about the impact of shadows and reflections on their sense of presence 17 expressed positive opinions. Of the 7 who said that the shadows and reflections had a negative impact, the common view was that this was because the reflection did not 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 may enhance realistic behavioral response. Of course the results here apply to a virtual environment that is designed to result in anxiety, and we do not know whether the results extend to more mundane applications. Moreover, although it is likely that it is the dynamic shadows and reflections causing the changed responses, we cannot disambiguate this from the general improvement in visual quality that results from recursive ray tracing –

though given the results in [7] it is unlikely that visual quality alone can account for the differences found here.

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

Acknowledgements

This research was carried out under the UK EPSRC Project 'Presence in a Virtual Light Field' (EP/C511824/1). Mel Slater's contribution was also supported through the EU 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] Sanchez-Vives, M.V. and M. Slater, *From Presence to Consciousness through Virtual Reality.* Nature Reviews Neuroscience, 2005. **6**(4): p. 332-339.

[2] Hendrix, C. and W. Barfield, *Presence within Virtual Environments as a Function of Visual Display Parameters.* Presence-Teleoperators and Virtual Environments, 1996. **5**(3): p. 274-289.

[3] Slater, M., M. Usoh, and Y. Chrysanthou, *The influence of dynamic shadows on presence in immersive virtual environments*, in *Selected papers of the Eurographics workshops on Virtual environments '95*. 1995, Springer-Verlag: Barcelona, Spain. p. 8-21.

[4] Mania, K. and A. Robinson. *The Effect of Quality of Rendering on User Lighting Impressions and Presence in Virtual Environments*. in ACM Siggraph International conference on Virtual Reality Continuum and its Applications in Industry. 2004. Singapore: ACM Press.200-205

[5] Zimmons, P. and A. Panter. *The Influence of Rendering Quality on Presence and Task Performance in a Virtual Environment*. in Virtual Reality, 2003. Proceedings. 2003: IEEE Computer Society.293-294

[6] Gibson, E.J. and R.D. Walk, *The Visual Cliff.* Scientific American, 1960. **202**(4): p. 64-72.

[7] Slater, M., M. Usoh, and A. Steed, *Taking steps: the influence of a walking technique on presence in virtual reality.* ACM Trans. Comput.-Hum. Interact., 1995. **2**(3): p. 201-219.

[8] Meehan, M., B. Insko, M. Whitton, and F.P. Brooks, *Physiological measures of presence in stressful virtual environments.* Acm Transactions on Graphics, 2002. **21**(3): p. 645-652.

[9] Khanna, P., I. Yu, J. Mortensen, and M. Slater, *Presence in response to dynamic visual realism: a preliminary report of an experiment study.* Proceedings of the ACM symposium on Virtual reality software and technology, 2006: p. 364-367.

[10] Usoh, M., E. Catena, S. Arman, and M. Slater, *Using presence questionnaires in reality.* Presence-Teleoperators and Virtual Environments, 2000. **9**(5): p. 497-503.

[11] Boucsein, W., *Electrodermal Activity*. 1992, New York: Plenum Press.

[12] Guger, C., G. Edlinger, R. Leeb, G. Pfurtscheller, A. Antley, M. Garau, A. Brogni, D. Friedman, and M. Slater. *Heart-Rate Variability and Event-Related ECG in Virtual Environments*. in Presence 2004: The 7th International Conference on Presence. 2004. Valencia, Spain: Universidad Politecnica de Valencia.240-245

Mel Slater is ICREA Research Professor at the Universitat Politécnica de Catalunya, Barcelona, Spain, and a Professor of Virtual Environments, University College London, UK. His research interests include global illumination algorithms for virtual reality and also scientific understanding of how people respond to virtual reality. He has a DSc from University of London and he can be contacted at mel.slater@icrea.es.

Centre de Realitat Virtual (CRV),
Edificio U,
Universitat Politècnica de Catalunya
C/ Llorens i Artigas 4-6
08028 Barcelona, Spain

Tel: +34 93 401 0761
Fax: +34 93 401 2592
Email: mel.slater@icrea.es

Pankaj Khanna currently works as a consultant in the financial services and maintains links to academia as an Honorary Research Fellow at University College London. He is also a PhD candidate with interests in global illumination, real-time physically-based rendering, and virtual reality.

Department of Computer Science
University College London
Gower Street
London WC1E 6BT, UK.

Tel: +44(0)7863206534
Fax: +44(0) 2073871397
Email: p.khanna@cs.ucl.ac.uk

Jesper Mortensen is a Graphics Programmer at Geomerics Ltd. working on real-time global illumination for the games industry. His research interest is rendering algorithms and particularly global illumination. He has an MSc in Vision, Imaging and Virtual Environments from University College London. He is a finishing PhD student on the topic of global illumination at University College London.
.
Geomerics Ltd.
City House
126-130 Hills Road
Cambridge
CB2 1RE
Tel: +44 (0)1223 450170
Fax: +44 (0)1223 361362
Email: jesper@geomerics.com

Insu Yu is a finishing PhD student at University College London. His interests include real-time global illumination on the GPU, especially based on spherical harmonics. He has an MSc in Vision, Imaging and Virtual Environments (VIVE) from the Department of Computer Science, University College London. He can be contacted at i.yu@cs.ucl.ac.uk.

Department of Computer Science
University College London
Gower Street
London WC1E 6BT, UK

Tel: +44 (0)20 7679 3876
Fax: +44 (0)20 7387 1397
Email: i.yu@cs.ucl.ac.uk