Virtual environments as research tools for environmental psychology: A study of the comparability of real and virtual environments

Yvonne A.W. Slangen- de Kort

Wijnand A. IJsselsteijn

Jolien Kooijman

Yvon Schuurmans

Eindhoven University of Technology

27 March 2001

Correspondence address:

Yvonne A. W, Slangen – de Kort

Faculty of Technology Management

Eindhoven University of Technology

PO Box 513

5600 MB Eindhoven

Netherlands

Virtual environments as research tools for environmental psychology: A study of the comparability of real and virtual environments

Virtual environments have the potential to become important new research tools in environment behaviour research. The present study is an exploration of the comparability of research findings in real and virtual environments. 101 participants explored an identical space, either in reality or in a computer-simulated environment. Additionally, the presence of plants in the space was manipulated, resulting in a 2 (Environment) x 2 (Plants) between Ss design. Employing a broad set of measurements we found mixed results. Performances on size estimations and a cognitive mapping task were significantly lower in the virtual environment. Factor analyses of bipolar adjectives indicated that although four dimensions were similar for both environments, a fifth dimension of environmental assessment – termed arousal – was absent in the virtual environment. In addition, we found significant differences on the scores of four of the scales. However, no significant interactions appeared between Environment and Plants. Experience of and behaviour in virtual environments has similarities to that in real environments, but important differences as well. We conclude that this is not only a necessary, but also a very interesting research subject for environmental psychology.

Virtual environments

Virtual Environments (VEs) have matured considerably in the past decade. Although there still is a diversity of technological problems to overcome, VEs are being applied in many areas such as entertainment, vehicle simulation, industrial and architectural design, training, medicine and other areas (s.f. Brooks, 1999). In this paper, we argue that the use of virtual environments as environmental simulations may provide a valuable tool for both research and practice in environmental psychology.

A VE is an artificial world, created with computers, which can give the observer a sense of 'being there' (presence) in the environment. For this, various input devices are needed to interact with or manipulate the environment (e.g. cursor keys, joystick, or head-tracker). The artificial world can be presented visually on a desktop display, a head-mounted display, or on one or more projection displays, sometimes combined with (spatialized) audio, haptic feedback and sometimes even scents or thermal cues (see Ellis, 1991 for an in-depth analysis of VEs). With the development of these types of media, the potential to provide viewers with an accurate representation of non-mediated experience has increased significantly. These mediated environments are thus able to provoke responses and behaviour similar to those portrayed in real environments (Lombard, 1995). The basis for this reasoning is the concept of *behavioural realism*, which is based on the premise that as a display better approximates the environment it represents, an observer's responses to stimuli within the display will tend to approximate those which he or she would exhibit in response to the environment itself (IJsselsteijn, de Ridder, Freeman & Avons, 2000; Freeman, Avons, Meddis, Pearson, & IJsselsteijn, 2000). To procure behavioural realism, one would have to present a sensorially rich and perceptually realistic environment that facilitates natural interaction between the user and the environment (as well as objects in it). Loomis, Blascovich and Beall (1999) argue that 'the ultimate representational system would allow the observer to interact "naturally" with objects and other individuals within a simulated environment or "world", an experience indistinguishable from "normal reality" ' (p. 557).

The use of simulations in research on environmental psychology

In the environment-behaviour domain, environmental simulations are undertaken for many purposes including (1) the training of environmental design students and planning professionals, (2) the assessment of people's preferences, behaviour patterns, and health when exposed to alternative environmental arrangements, (3) the visualization of complex settings prior to their design and construction, and (4) the incorporation of observers' evaluations of simulated settings into the planning and design of new environments or the renovation of existing ones (Stokols, 1993). In considering these different uses, it is hardly surprising that the term 'environmental simulation' covers a wide range of different types of representations of our natural and human-made habitat. An early attempt to structure this broad range was McKechnie's (1977) typology of simulation techniques describing two dimensions on which simulations vary (abstract-conceptual vs. concrete-perceptual and static vs. dynamic). More recently, Ozel (1993) proposed a classification for environmental representations relevant for the use of computers in the world of modelling and simulation, distinguishing between (1) visual representations of objects, (2) systemic representations of objects with their configurational and nonconfigurational attributes, and (3) representations of systems, processes and events in architectural environments. The differences between these two typologies only emphasize the diversity in this domain and indicate the frequent use scientists and environmental psychologists and designers make of simulations of the physical environment. All these simulations have in common that they attempt to represent some aspects of the environment as accurately as possible to assess human responses to them.

In the history of environment behaviour studies the most extensive effort to validate an environmental simulation was the Berkeley Environmental Simulation Project (Appleyard & Craik, 1978; Bosselman & Craik, 1987). Few other studies offer comparisons between physical models and the real setting (e.g. Hunt, 1984; Kaplan, Kaplan & Deardorff, 1974; Seaton & Collins, 1972). In reviewing the material, Kaplan (1993) concludes that there is insufficient empirical evidence to conclude that simulations elicit identical reactions to those in real–world settings. In the past, the greatest emphasis has been on static modes of simulation. In their review, Craik and Feimer (1987) conclude that the small number of systematic studies with dynamic simulations suggest that descriptive and evaluative responses are comparable to those obtained for direct, on-site presentations. However, evidence concerning cognitive and behavioural responses is less clear. Based on a meta-analysis of 11 papers, Stamps (1990) does conclude that there are strong relationships between preferences obtained in the environment and preferences obtained through photographs. Besides this, others have shown that viewing (nature) pictures has physiological effects which are similar to the experience of natural environments (Ulrich et al., 1991).

One of the major problems in the validation of simulations is the fact that progress in environmental psychology has not yielded a comprehensive array of standard response measures (Bosselman & Craik, 1987). The Berkeley Project used numerous response formats; researchers focused on the 'equivalence of response' to the different ways the environment was experienced as a critical validation issue. Other studies comparing physical models with real settings (e.g. Hunt, 1984; Kaplan, Kaplan & Deardorff, 1974; Seaton & Collins, 1972) use a quite different choice of dependent variables. The environmental assessment paradigm consists of a wide range of cognitive, affective, and behavioural responses (Craik & Feimer, 1987). At present there is no standard set of responses, although some measurement instruments are used more often than others. For instance, the use of semantic differentials and sketch maps is quite common (Canter, 1991).

The utility and validity of virtual environments as simulations

The utility of any simulation type or technique depends on the context in which it is used, but VEs hold important qualities for use as simulations of the physical environment in diverse areas: they are dynamic and perceptually rich, sometimes even multi-modal, 3-dimensional and interactive. It is argued that in the near future, VR will approximate reality to such a degree that we will frequently conduct research in VEs rather than in real environments. The use of VEs can increase the power of experimental research, by providing both high ecological validity and experimental control, given that they show an appropriate level of experiential and behavioural realism (Loomis et al., 1999). Therefore, VEs are believed to have the potential of becoming an important research tool for psychological research. If this is true, this should certainly hold for the subdomain in which the interaction between individuals and their physical environment is the central object of study: environmental psychology.

Even though the technique has been developing at a steady pace towards a more realistic experience, there still are a few major shortcomings that lead to differences in perception between real and virtual environments. For instance, perception in many VEs is uni-modal: only the visual senses are stimulated. Aspects of the VE itself, like the resolution of the projection and the level of detail, also differ from the real environment. Because the physical body is absent, there is no correspondence between what you see and the movement and position of the body (although sometimes a virtual body is generated to overcome this limitation). Another important difference is that technical devices are necessary to navigate through a VE, which can make the experience even more unnatural. All of these aspects contribute to the fact that the experience of a virtual environment differs from the real environment.

Whether research in VEs will – to a smaller or larger degree – substitute for research in the real world remains to be seen and will definitely require significant progress in technology and a more thorough understanding of the human factors issues involved (with complex VEs, end-to-end system latency is still the major technical limitation (Brooks, 1999)). However, the fact that VR-technology has already been embraced by large numbers of professionals in design, urgently calls for research to increase our understanding of person-environment transactions in virtual worlds. The need for more research that addresses applications of perceptual simulations in general and related questions of validity and reliability has been stressed ever since the emergence of environmental simulation as a research paradigm. However, the number of validation studies has been scarce (Bosselman & Craik, 1987). Zube and Simcox (1993) argue that the risk of 'naive' acceptance of new, high-tech simulation technologies is even bigger: '*there is little evidence that the necessary psychological research has been or is being carried out. To the contrary, the literature suggests that the development of* *perceptual simulation technologies is often advancing under the assumption that new technologies are de-facto meeting criteria of validity and reality, regardless of the area of application*' (p. 275).

Rationale

The present study is a first general exploration of the comparability of research findings in real and virtual environments within the domain of environmental psychology. Some related research was found in other scientific domains. For instance, in discussing the literature on filmed and televised presentations, Lombard (1995) states that there is some compelling empirical evidence that media users react to televised presentations in some of the same ways that they react to non-mediated events, objects, and people despite the fact that mediated presentations provide a limited reproduction of the non-mediated experience. He suggests that these responses should be termed direct, since the users were responding to the actual events and objects rather than to mere portrayals of events. In his own research he even extended these findings to the realm of social interaction, where he studied direct emotional responses to people on the screen as a function of interpersonal distance (Lombard, 1995). Little has been done on environmental perception and assessment in VEs, apart from studies focusing on basic task performance (e.g. discrimination and recognition, tracking, and distance judging, for a review see Nash, Edwards, Thompson, & Barfield, 2000). One qualitative study by Murray, Bowers, West, Pettifer and Gibson (2000) suggested a continuous relationship between real and virtual worlds based on an analysis of the interaction of people with a virtual city. Participants were seen to attribute real-world properties and expectations to the contents of the virtual world. However, besides reported similarities between non-mediated and mediated events, and between real and virtual environments, important differences are also reported, e.g. navigation within VEs is not as simple as in the real world and can lead to a higher tendency of a user becoming lost (Witmer, Bailey, Knerr, & Parsons, 1996).

The basis for the evaluation of realism in a simulation is a comparison of responses to real places and simulated places (Bosselman, 1993). In the present study we aimed to compare measurements in a real-world environment with measurements in an accurate computer simulation of that environment. Since this was an exploratory study and for want of a standard set of measurements, we employed a broad and fairly general set of measurements including (1) estimations of sizes and heights, (2) a cognitive mapping task, (3) subjective evaluations of the environment using bipolar adjectives, and (4) qualitative analyses of perceived (behavioural) functions afforded by the environment. This set of measurements thus consisted of both evaluative and performance measures and tried to assess cognitive, affective and behavioural responses. In addition to comparing the two environments directly, we set out to compare the influence of a subtle change in these environments, i.e. the presence or absence of plants. Plants were chosen to serve as manipulation, because these were both available in the real environment and technically manageable in the VE. The expectation was that participants would evaluate both types of environment more positively for the situation with plants than without plants. We expected to find some differences in scores between the real world and the virtual world, because the real world is always richer with stimulation and because interaction with it is more natural than interaction with a virtual environment. However, we did not expect that this would cause differences in the effects of the additional presence-of-plants manipulation for both worlds. In other words, we did not expect to find interaction effects.

Purposive evaluation

It is hard to evaluate the relevance of various findings in a study in absence of a practical context. As Canter (1991) stated, place evaluation only makes sense when viewed in relation to the purposes for which the place is being evaluated. In a relatively context-free study like the present one, there is no purposive evaluation, which not only makes it harder to formulate sensible measurements, but also makes it impossible to weigh the pros and cons in pursuit of a ‘go-no-go’ decision regarding the use of this technology as a simulation method in environmental design. However, we felt that as a first exploration of the utility of this technique, a general, context-free study was warranted. In order to give participants a goal during exploration of the space and filling out the questionnaire, we therefore employed a rather general instruction to assess multifunctionality. The idea was that this would at least incite them to go beyond the purely aesthetic characteristics of the building and also assess its content.

Method

Participants and Design

A total of 101 participants (62 male and 39 female students and recent graduates of various disciplines) were recruited for the experiment via e-mail and poster announcements at the Technische Universiteit Eindhoven. Respondents' ages ranged between 18 and 35 years (M = 23.1). Every participant received 7.50 Dutch guilders for their co-operation (equivalent to $ 4). The participants were randomly distributed over the four experimental conditions following a 2 (Environment: real (RE) vs. virtual (VE)) x 2 (Plants: yes vs. no) between-subjects design.

Materials

The environment in the study was the main hall of the Multimedia Pavilion, a building on the premises of the university. This fairly new building houses several small companies and has an interesting and colourful design and interior. It also houses the offices and demonstration room of Calibre BV, a company founded by the Technische Universiteit Eindhoven that is specialised in 3D computer graphics and computer simulations. This company also simulated their own working environment, resulting in an accurate virtual environment of the Multimedia Pavilion. Figure 1 shows similar views into the building taken in the RE and the VE respectively. Calibre BV supplied the equipment and the simulation of the environment, as listed in Table 1.

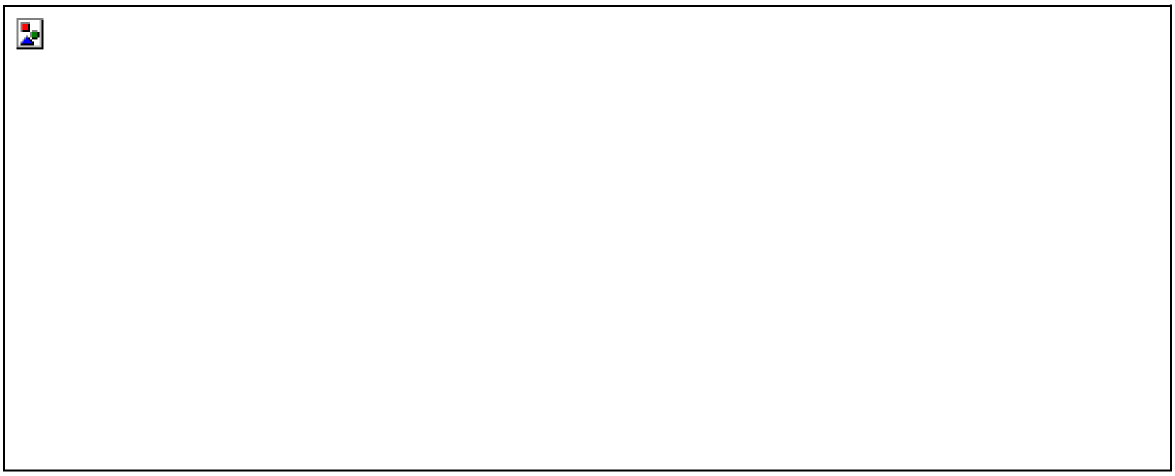


Figure 1: Real and Virtual Environment

Note: left picture was taken at daytime, but the study was performed in the evening, with similar types and levels of lighting for RE and VE

Table 1: Technical specifications of the materials used

Hardware	Computer	Silicon Graphics Onyx 2 Infinite Reality Station 2x MIPS R10000 processors, 1GB intern memory 2 Geometry Engines, 1 Raster manager
	Screen	Back projection with Barco 1208 data projector Screen: 2.40 x 1.70 m
	Periphery	BG-systems Flybox, 12 button professional serial joystick , 3 degrees of freedom
Software		dvMockup simulation software from PTC-Division With plugins and adjustments by Calibre

Procedure

Participants were received in a building adjacent to the Multimedia Pavilion and were taken to the Pavilion by a research assistant. There, they were guided either to the main hall or the demonstration room in this same building, depending on the experimental condition (RE vs. VE respectively). Participants in the VE condition did not see the main hall in reality. The experiment was conducted in the evening, so there would be no people working in the offices in the same building and lighting conditions could be controlled.

Participants explored the main hall of the Multimedia Pavilion, either in the RE or the VE, and either with or without plants. All participants were told the experiment was about multifunctional environments and were asked to explore the space very carefully. They were asked to imagine what functions the space could have, to give them a goal during the exploration. The starting point was the same for both RE and VE. After 3.5 minutes participants were given a signal whereupon they returned to the starting point. Subsequently they filled out a questionnaire.

Prior to their exploration participants in the VE received a short instruction to help them navigate in a (different) VE. They were located at a distance of 1.5 meters from the screen and the eye-to-floor distance was set at 1.65 meter for every participant. The horizontal visual angle was 77°. Participants in the VE were specifically asked not to judge the simulation, but the environment.

The questionnaire consisted of five parts (in the order as they appeared in the study):

(1) basic descriptive data, consisting of age, gender, study background and previous experience with the Multimedia Pavilion;

(2) a semantic differential scale, which consisted of 29 bipolar adjective 6-point items, corresponding to those generally used in environmental assessment research (items are listed in Table 3);

(3) height and colour estimations, where participants were asked to estimate the heights of the room and the doors, both with a margin of 10 cm, and to choose the colours of the chequered floor from a sample of 13 colours;

(4) functions of the space, where they listed the functions they thought the space could provide ('afford') its users in an open response format (the researchers specified no minimum or maximum number); and

(5) sketch maps. In this phase participants were asked to draw a map of the environment they had explored as precisely and correctly as possible. The sketches were evaluated afterwards by three raters on five 5-point scales: completeness of sectioning of the space, correctness of relative proportions, number of omissions and augmentations, correctness of curved and sloping parts, and a general score for the sketch as a whole. The inter-rater reliability for these five items was analysed using Kendall's W, which showed that all values were satisfactory (ranging from .73 to .87). Averaged scores of the three observers' ratings are used in subsequent analyses.

The experiment lasted approximately 30 minutes.

Results

The primary statistical procedures used to analyse the data were analysis of variance (ANOVAs) and principal components factor analysis. In this section we will first report the results relating to the performance measures, then those related to the evaluative measures. First ANOVAs were conducted to test the differences with respect to the deviation between the estimated and actual heights of the doors and the room. Next, the five scores for the quality of participants' sketch maps were factor analysed, combined into one average score and tested with an ANOVA. Then factor analyses of the bipolar adjective items were used to explore connotative dimensions for the virtual and real conditions separately. Although this does not provide a hard or solid test, the comparison of these factor structures could give a first indication of the equivalence of both environments. Subsequently, the scores on five dimensions were computed and analysed with separate ANOVAs according to the full model. The section closes with the results regarding the functions of the space.

Estimations

Participants made estimations of the height of the room and the doors. These are shown in Figure 2. Differences were tested with two additional full-model ANOVAs with door height and room height as dependent variables. Both scores were significantly different for the two levels of Environment ($F(1,97)=4.26$; $p=.042$ and $F(1,96)=35.18$; $p<.001$ respectively). Participants in the RE made estimations that were more accurate than those in the VE. Participants in the VE tended to underestimate the heights of the doors and the room, which are unusually high in this building. There was no significant main effect of Plants. However, for the scores of door height, a significant interaction appeared, $F(1,97)=3.98$; $p=.049$. This interaction reflected no effect of Environment in the 'with plants' condition but a strong effect of Environment in the 'without plants' condition. Possibly, the plants served as a familiar size cue for the estimation of door heights in the VE.

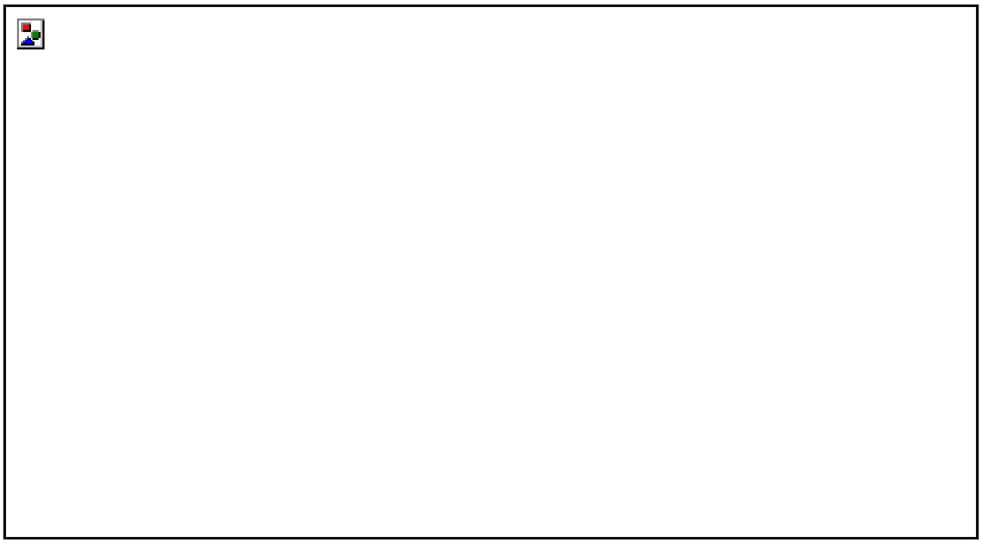


Figure 2: Scores on the height estimations of doors and room for both levels of Environment and Plants

Participants also had to judge the colours of the chequered floor: they chose 2 colours from 13 samples. Scores were computed based on the correctness of the colours participants chose. They received 2 points for choosing one of the exact colours, 1 point for a colour that came close to the exact colour, 0 points for the other colours. Participants' scores (ranging between 0 and 4) were analysed with a Mann-Whitney U-test, which indicated that there was no significant difference regarding this measure for the different media (U=1137.5, N=100 and p=.420). In other words, participants in the RE did not perform better than participants in the VE.

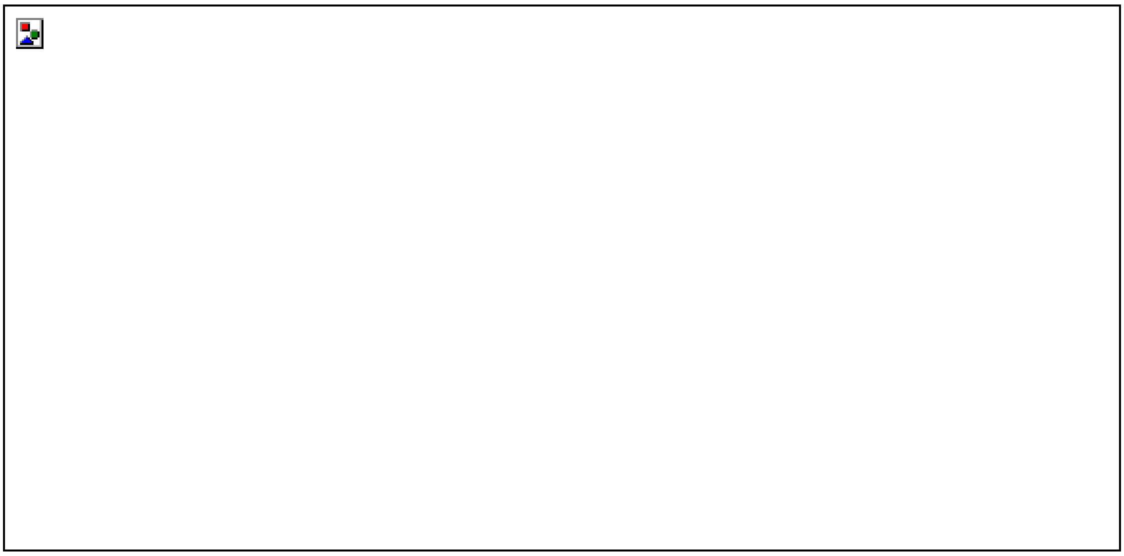
Sketch maps

Participants' sketch maps of the environment were assessed on 5 scales – completeness of sectioning of the space, correctness of relative proportions, number of omissions and augmentations, correctness of curved and sloping parts, and a general score for the sketch as a whole. A factor analysis was performed which indicated that these five variables could be reduced to one variable. The reliability of this scale was high, Cronbach's alpha = 0.89. The means and standard deviations are shown in Table 2 and Figure 3.

Univariate Analysis of Variance using Environment (real/virtual) as independent variable and the average score for the sketches as dependent variable revealed a significant difference between the RE and the VE, F(1,99)=7.809, p = .006, indicating that the sketch maps of participants in the RE were more complete and correct than those of participants in the VE.

Table 2: Quality of sketch maps – scores on separate scales and average

	Separate scales					**Average**	
Environment	sections	proportions	omissions	curves	general	M	(SD)
real	3.31	2.22	2.26	1.50	1.95	2.25	(.73)
virtual	2.83	1.84	1.79	1.28	1.53	1.86	(.70)
Total	3.07	2.03	2.02	1.38	1.75	2.05	(.74)



Note: scores range from 1 to 5

Figure 3: Scores on the scales judging participants' sketch maps for both environments

Bipolar adjectives

Factor analyses with Varimax rotation of the 29 bipolar adjective items were used to develop indexes for the complete sample and for the VE and RE conditions separately. A five-factor solution was chosen in the RE. We termed these factors Evaluation, Ambience, Arousal, Privacy, and Security. In the VE, a four-factor solution was chosen. Component loadings for the two types of environment separately are reported in Table 3. Three factors were almost identical for RE and VE: Evaluation, Privacy, and Security. The fourth factor in the VE was similar to the Ambience factor in the RE, but consisted of two additional items (calm-busy and monotone-varied). In the RE these items were clustered with items related to symmetry and legibility of the space and with the item chaotic-ordered. This did not occur in the VE. Forcing a five-factor solution for the VE (or a four-factor solution for the RE) did not result in better interpretable data.

Table 3: Bipolar adjective items: component loadings after Varimax rotation and scale reliabilities

	Real environment					**Virtual environment**			
	1	2	3	4	5	1	2	3	4
Evaluation	**.92**					**.89**			
ugly-beautiful	**.88**					**.85**			
tasteless - tasteful	**.79**					**.69**			
unpleasant - pleasant	**.77**					**.79**			
boring - interesting	**.72**	.45				**.75**			
unattractive - attractive	**.65**	.44				**.64**			.40
meaningless - impressive	**.63**	.43				**.75**			
uninviting - inviting	**.49**	.43			.47	**.58**			
artificial - natural	**.47**	.37				**.58**			.36
gloomy - cheerful	**.61**	.43				.39	**.54**		
Ambience		**.84**					**.82**		
impersonal - personal		**.74**				.39	**.73**		
cold – warm	.16	**.74**				**.65**	.54		
grey - colourful		**.72**				.41	**.51**		
business-like - playful		**.71**					**.68**		
not cosy - cosy	.39	**.68**				.46	**.54**		
bare - decorated	.36	**.55**				.36	**.50**		
Arousal			**.77**				**.56**		
chaotic - ordered			**.87**				.44		**-.59**
asymmetrical - symmetrical			**.72**					**-.49**	
illegible – legible			**.71**					**-.74**	-.39
calm - busy			**.55**		-.48		**.77**		
monotone - varied	.37	.40	**.49**			.46	**.54**		
Privacy				**.73**				**.53**	
public - private				**.82**				**.59**	-.40
common - individual				**.74**			-.40	**.58**	
open - enclosed	-.39			**.65**				**.36**	
light - dark	-.48			**.57**				**.46**	
Security					**.72**				**.66**
scary - relaxing	.47				**.70**				**.68**
unsafe - safe		.37			**.65**				**.66**
inaccessible - accessible				-.44	**.59**				**.59**
threatening - protecting		**.53**			.26		.40		**.66**

Not included					
quiet - lively	.45	**.50**	-.43		**.61**

Note: Results are reported for RE and VE groups separately. All component loadings >.35 are reported. Highest loadings for every analysis are printed in bold. Scale reliability: Cronbach's alpha is reported for every scale. The item quiet-lively was not included in any of the scales since it decreased scale reliability.

It was decided to compute the scales for all the data based on the five-factor solution in the RE. This solution was clear and based on assessment in an actual environment. Reliability analyses were performed for the complete sample. This resulted in satisfactory to high reliabilities for the complete sample (Cronbach's alpha .64 - .91). The alpha values for the five scales for both environments separately are reported in Table 3. Reliabilities in the VE are somewhat lower than in the RE. Nevertheless, we computed five identical scales for both environments, so that the scores could be compared.

Next, the scores on these five scales were analysed with separate ANOVAs using Environment and Plants as independent variables. This resulted in modest but significant effects of Environment on Evaluation, $F(1,95)=9.51$, $p=.003$; Ambience, $F(1,96)=8.05$, $p=.006$; Privacy, $F(1, 96)=5.08$, $p=.026$; and Security $F(1,97)=4.07$, $p=.046$. All scores were higher in the RE condition (see Figure 4); means and standard deviations are reported in Table 4. The effect of Plants only yielded a marginally significant effect for Privacy, $F(1,97)=3.61$, $p=.06$, resulting in higher scores (more privacy) in the no-plants condition. No significant interactions between Environment and Plants appeared.

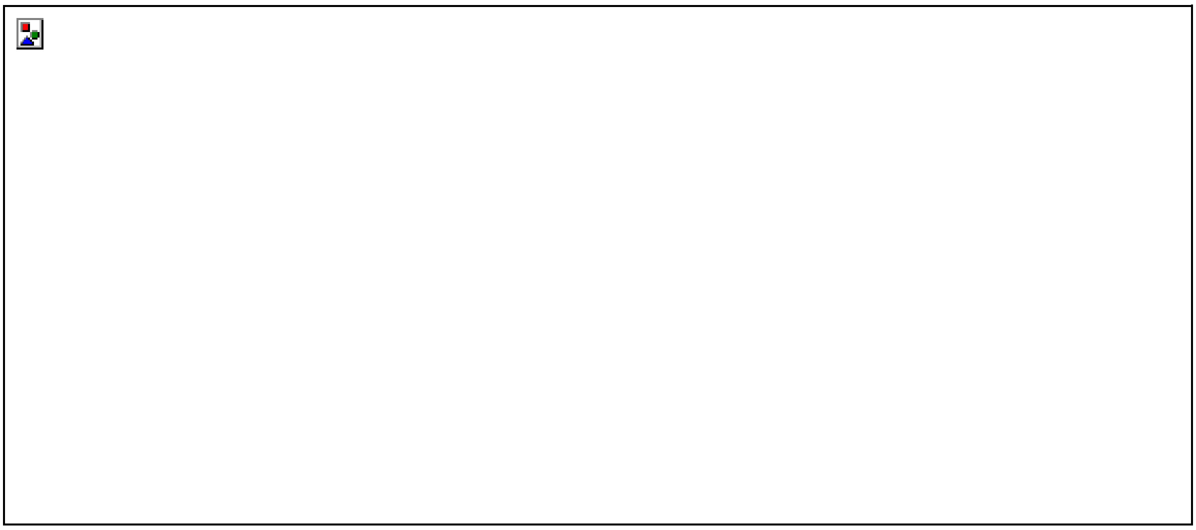


Figure 4: Scores on the five evaluative scales for both levels of Environment and of Plants

Table 4: Scores on the five evaluative scales for both levels of Environment and Plants

		Plants			
		with plants		without plants	
	Environment	Mean	(SD)	Mean	(SD)
Evaluation	real	4.02	(.97)	4.19	(.90)
	virtual	3.66	(.78)	3.44	(.90)
Ambience	real	3.39	(.97)	3.62	(.86)
	virtual	3.19	(.91)	2.81	(.81)
Arousal	real	3.10	(1.04)	3.26	(.94)
	virtual	2.99	(.87)	2.87	(.82)
Privacy	real	2.52	(.82)	2.77	(1.04)
	virtual	2.10	(.67)	2.46	(.72)
Security	real	4.48	(.69)	4.33	(.73)
	virtual	4.15	(.57)	4.08	(.85)

Note: Scales ranging from 1 to 6; scores in parentheses are standard deviations

Afforded functions

Participants listed the functions they thought the space could afford. These functions were categorized and counted by an independent rater. The categories and their frequencies are reported in Table 5. Participants in the RE condition listed significantly more functions ($\underline{M}$=3.2) than participants in the VE condition ($\underline{M}$=2.7) did, $\underline{F}(1,99)=5.15$, $\underline{p}=.025$. Chi-square values were calculated and resulted in significant differences between the real and VE for functions related to offices and education – participants reported these functions more often in the VE – and for functions related to social gatheringss – these were reported more frequently in the RE. No significant differences were found between the with and without plants conditions.

Table 5: Frequency of afforded functions of the space as perceived and listed by participants

	Environment			
	real	virtual	**Chi sqr**	**Sig**
Function	N = 49	N = 52		
Coffee or lunch breaks	27	20	2.81	n.s.
Passage/ corridor	22	22	.21	n.s.
Reception/ meeting place	20	19	.06	n.s.
Socials gatherings	**27**	**9**	22.99	<.001
Offices	**8**	**26**	12.81	<.001
Waiting	15	8	3.33	n.s.
Education	**2**	**16**	12.27	<.001
Expositions	7	2	-	

Public/municipal space	1	7	-
Foyer theatre	4	1	-
Library	1	2	-
Hospital	2	1	-
Apartments	2	1	-

Note: Chi square was not computed for functions with total N<10.

Discussion

The goal of the present study was to explore the comparability of research findings in real and virtual environments (RE vs. VE) within the domain of environmental psychology. Employing a broad set of measurements we found mixed results. Performance measures (height estimations and a cognitive-mapping task) were significantly lower in the VE than in the RE. Differences also appeared for evaluative measures. Factor analyses on bipolar adjective items in both environments separately indicated that four dimensions were very similar, but a fifth dimension was missing in the VE. In addition, we found modest but significant differences between the RE and the VE on the scores on four of five computed scales. However, there were no significant interactions between Environment and Plants, the second manipulation of the environment.

Performance measures: estimations and sketch maps

Participants in the RE performed significantly better than the ones in the VE: they made height estimations that were more accurate and their sketch maps were more complete and correct. Several reasons can be given to explain these findings. One possibility is that participants in the simulation performed worse on the cognitive mapping task owing to the cognitive load posed by the medium of Virtual Reality itself and operating the joystick for navigation. However, we feel that this load was only minimal and could hardly explain the remaining differences regarding the height estimations. A second possibility is that these differences are attributable to the relatively small field of view. The field of view – the visual angle of the image – in the VE was smaller than the regular human field of view. The relevance of this variable in various search and estimation tasks in virtual reality has been demonstrated in earlier studies (e.g. Venturino and Wells, 1990; for a review see Nash et al., 2000). Venturino and Wells argue that a larger field of view allows the user to integrate larger amounts of spatial information more quickly. Thirdly, the unnaturalness of navigating through a VE – using a joystick instead of actually walking and the absence of (a representation of) one's body – may make it more difficult to infer spatial information and create an accurate overall image of the environment. We conclude that these findings call for additional research, in which the role of these possible determinants is investigated.

The interaction between Environment and Plants on the estimation of door height indicated that perhaps the plants served as a familiar-size cue in height perception for participants in the VE. This effect did not appear in the RE where in fact participants performed worse in the with-plants condition. People in the RE probably used their own height as a cue, whereas participants in the simulation did not (although theoretically they could have since eye-height was established at the average human eye-to-floor distance). The findings are somewhat ambiguous, especially since the plants did not serve as a cue in the estimation of the height of the room, but perhaps the size of the plants was too small (relative to the height of the space) for that purpose.

Evaluation: bipolar adjectives and afforded functions

Evaluation of the environment as measured with 29 bipolar adjective items resulted in five connotative dimensions in the RE. We termed these factors Evaluation, Ambience, Arousal, Privacy, and Security. Only four dimensions appeared in the VE, the Arousal dimension was absent there. Some of the items clustered with Ambience, the others with Privacy and Security. This finding may reflect a difference that was coincidental or specific for this simulation, but could also reflect a difference in perception and experience between real and virtual environments in general. In the RE items related to arousal (calm-busy, monotone-varied) clustered with items related to space coherence and legibility (illegible-legible, asymmetrical-symmetrical, chaotic-ordered). The fact that this did not happen in the VE could be related to the findings that were reported in the former section. There we argued that inferring spatial information and constructing an overall image of the environment may be more difficult or at least different in VEs as compared to REs. It is interesting to pursue these issues in future investigations.

In spite of these differences in factor structures, five identical scales were computed for both environments. For these dimensions, significant differences were found between the two environments on all scales except the Arousal scale. Participants evaluated the RE more positively on Evaluation, Ambience, Privacy, and Security. The difference is not likely to be attributable to differences in colour perception, since both groups scored similarly on the floor-tile colour test.

Differences could be due to the lower level of experiential realism in the VE and the quality and vividness of the simulation, which still is lower than that of real environments in spite of its sophistication. This is mainly due to a lower level of detail, the absence of reflective surfaces, and the resolution of the projection. The differences in the scores for these dimensions were also reflected in the type of functions participants said the space afforded: in line with higher scores on Evaluation, Ambience, and Privacy, social gatherings (receptions and socials) were mentioned more frequently in the RE whereas more formal activities (offices and education) were mentioned less frequently there. Participants assessing the VE also scored lower on subjective security, which could be due to the lower levels at which the space could be overviewed.

For research purposes, modest differences in scores between REs and VEs do not necessarily pose a problem. Two aspects are more relevant in this respect. Firstly, it is imperative that effects of manipulations in real and simulated environments are equal, as was the case in the present study. No significant differences arose between the two media with regard to the presence or absence of plants, i.e. there were no interactions between the factors Environment and Plants. However, the plants-manipulation showed only one marginally significant effect on the Privacy scale. The manipulation may have been too subtle, although we heard numerous distinctly positive remarks from the people normally occupying the adjacent office spaces. We therefore recommend additional research with stronger or other manipulations of both environments.

Second, it is relevant to determine the similarity of the nature of environmental perception and experience as for instance reflected in the factor structure of the bipolar adjective items. The fact that four out of five dimensions were very similar illustrates that the use of VEs as prototypes or simulations looks promising. However, a similarity in factor structure across media conditions itself grants no assurance that specific research sites will be described or evaluated comparably (Craik & Feimer, 1987). Furthermore, it should be noted that conclusions based on factor analyses with only one type of environment are somewhat premature. Strong conclusions should only be drawn after analysing a broader set of environments, both in reality and virtual reality.

Advantages of the use of virtual environments in environmental psychology

In the preceding sections we reported both similarities and differences in experiences for real and virtual environments. However, a number of favourable characteristics of the use of VEs as simulations have not yet been elaborated. For instance, Catalano and Arenstein (1993) pointed out the danger of the use of purely visual, high-quality images as simulations, for instance in community development processes. They argue that an important problem is that new simulation technologies, giving more weight to the 'profound sensation of aesthetic effects' may involuntarily increase the weight people place on these effects in comparison to other outcomes. Catalano and Arenstein state that the literature on social traps is replete with examples of seemingly poor judgment attributable to the fact that sensory effects are often weighted more heavily than adverse outcomes that are understood but not experienced. This would plead for simulations that perhaps are somewhat less vivid and rich than reality with regard to visual sensations, but which afford different types of experience with the future environment. By virtue of its characteristics, a VE affords more possibilities for meaningful interaction and walk-through experiences than many other types of simulations. This would offer greater attention to for instance contextual qualities, social amenities, accessibility and functionality.

VEs may increase the ease with which research can be done, and may offer more possibilities for providing both ecological validity and experimental control (for a discussion see Loomis et al., 1999) Besides this, they can also give researchers the ability to perform manipulations that are difficult or impossible to perform in reality. An example of the contribution virtual reality can make to research was given by Rose and Foreman (1999), regarding the role of active interaction with an environment rather than its passive observation in terms of cognitive processing. It is now possible to present rich and vivid 'worlds' to research participants and to carefully manipulate variables in this world, or in participants' interaction with it, for instance active vs. passive navigation, while keeping all other variables constant. In addition, the technology provides many possibilities for measuring and recording relevant behaviours and interactions of people with the environments with relative ease (e.g. eye-tracking or behaviour mapping).

However, before we can use VEs as research tools, validity and comparability of experience and behaviour in VEs in itself are important research subjects for environmental psychology. Interesting questions can be asked and are discussed in research domains outside the perspective of environmental psychologists. For instance, Murray et al. (2000) write: 'Would people's methods of exploration [in VEs] be analogous to how people move around in real environments? If not, how and why do they differ?' (p. 436). This just goes to illustrate that there is a strong and rapidly developing research domain on virtual reality and VEs. It is asking many of the questions environmental psychologists could help to answer and moreover, is also producing highly relevant theories and research data for our domain. We would like to argue that it would be interesting to share ideas, methodologies, and important concepts between these fields. One such important concept is 'presence'.

Presence

The sense of presence is the defining experience for virtual reality (Steuer, 1992). Lombard & Ditton (1997) define presence as the 'perceptual illusion of non-mediation', i.e. the extent to which the person fails to perceive or acknowledge the existence of a medium during a technologically-mediated experience. This 'experience of non-mediation' is probably related to lower levels of cognitive load for navigation purposes and the medium itself. Secondly, it is a strong indicator of 'experiential realism', which could be an important determinant of the comparability of subjective evaluations between real and virtual environments (IJsselsteijn et al., 2000). Numerous researchers are now trying to construct a solid and valid scale for measuring presence and are studying its multidimensional structure. The relevance of presence in the context of environmental psychology research with VEs does however call for a measuring instrument that could be used in both real and virtual environments. The authors are currently in the process of developing such an instrument.

Conclusion

The experience in a VE is different from that in a RE in several ways. The main differences reported in the present study in this respect are those related to the problems in integrating spatial information to configurational knowledge of the space. The present study also showed modest but significant differences in evaluations between a virtual and the real environment. Earlier studies with other types of simulations have shown similar findings between the real world and simulations or between several types of simulations (s.f. Bosselman & Craik, 1987; Feimer, 1984; Kaplan, 1993). Although these differences can be highly relevant in simulations that function as a tool in the (participatory) design process, the focus for scientific research generally is on structural similarity of experience and equality of effects of manipulations between the real world and the simulation.

In spite of the reported differences, there probably are numerous situations in which VEs better approximate perception of and interaction with spatial and architectural features of naturally occurring settings than other simulation types (e.g. drawings, scale models, or slides) and offer ample means to carefully register and study behaviour in environments. We conclude that it is interesting to pursue this type of research, not only because of the possible benefits of the future use of VEs as research instruments, but also because VEs and people's experience of and interaction with them are interesting research objects in themselves. We argue that VEs could prove a valuable simulation technique for a broad range of scientific and applied projects once we have discovered its basic requirements for valid research findings.

Acknowledgements

The authors wish to thank Calibre BV for their support, supplying the simulation and placing the hard- and software at their disposal.

References

Appleyard, D. & Craik, K. H. (1978). The Berkeley Environmental Simulation Laboratory and its research programme. *International Review of Applied Psychology, 27*, 53-55.

Bosselman, P. (1993). Dynamic simulations of urban environments. In R. W. Marans & D. Stokols (Eds.), *Environmental Simulation; Research and Policy Issues* (pp. 279-302). New York: Plenum Press.

Bosselman, P. & Craik, K. H. (1987). Perceptual simulations of environments. In R. B. Bechtel, R. W. Marans, & W. Michelson (Eds.), *Methods in environmental and behavioral research* (pp. 162-190). New York: Van Nostrand Reinhold Company.

Brooks, F. P. (1999). What's real about virtual reality. *IEEE Computer Graphics and Applications*, 19 (6), 16-27.

Canter, D. (1991). Understanding, assessing, and acting in places: Is an integrative framework possible? In T. Gärling & G. W. Evans (Eds.), *Environmental cognition and action: An integrated approach* (pp. 191-209). Oxford: Oxford University Press.

Catalano, R. & Arenstein, W. (1993). Strategies of environmental simulation; Theoretical, methodological, and policy issues. In R. W. Marans & D. Stokols (Eds.), *Environmental Simulation; Research and Policy Issues* (pp. 303-313). New York: Plenum Press.

Craik, K. H., & Feimer, N. R. (1987). Environmental assessment. In D. Stokols & I. Altman (Eds.), *Handbook of Environmental Psychology* (pp.891-918). Chichester: Wiley-Interscience.

Ellis, S. R. (1991). Nature and origin of virtual environments: a bibliographical essay. *Computer Systems in Engineering*, 2, 321-346.

Feimer, N. R. (1984). Environmental perception: the effects of media, evaluative context, and observer sample. *Journal of Environmental Psychology*, 4, 61-80.

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

Hunt, M. E. (1984). Environmental knowing without being there. *Environment and Behavior, 16*, 307-334.

IJsselsteijn, W. A., de Ridder, H., Freeman, J. & Avons, S. E. (2000). Presence: Concept, determinants and measurement. *Proceedings of the SPIE* 3959, 520-529.

Kaplan, R. (1993). Physical models in decision making for design; Theoretical and methodological issues. In R. W. Marans & D. Stokols (Eds.), *Environmental Simulation; Research and Policy Issues* (pp. 61-86). New York: Plenum Press.

Kaplan, R., Kaplan, S., & Deardorff, H. L. (1974). The perception and evaluation of a simulated environment. *Man-Environment Systems, 4*, 191-192.

Lombard, M. (1995). Direct responses to people on the screen: Television and personal space. *Communication Research*, Vol. 22 (3), 288-324.

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

Loomis, J. M., Blascovich, J. J. & Beall, A.C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behaviour Research Methods, Instruments, & Computers 1999, 31 (4), 557-564.*

McKechnie, G. E. (1977). Simulation techniques in environmental psychology. In D. Stokols (Ed.), *Perspectives on Environment and Behavior: Theory, Research, and applications* (pp. 169-189). New York: Plenum Press.

Murray, C.D., Bowers, J.M., West, A., Pettifer, S., & Gibson, S. (2000). Navigation, wayfinding, and place experience within a virtual city. *Presence, 9* (5), 435-447.

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A review of presence and performance in virtual environments. *International Journal of human-computer Interaction, 12*, 1-41.

Ozel, F. (1993). Computer simulation of behavior in spaces. In R. W. Marans & D. Stokols (Eds.), *Environmental Simulation; Research and Policy Issues* (pp. 191-212). New York: Plenum Press.

Rose, D. & Foreman, N. (1999). Virtual reality. *The Psychologist*, 12 (11), 550-554.

Seaton, R. W., & Collins, J. B. (1972). Validity and reliability of ratings of simulated buildings. In W. J. Mitchell (Ed.), *Environmental Design: Research and Practice* (pp. 6-10-1 to 6-10-12). Washington, D.C.: EDRA.

Stamps, A. E. (1990). Use of photographs to simulate environments: A meta-analysis. *Perceptual and Motor Skills, 71*, 907-913.

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

Stokols, D. (1993). Strategies of environmental simulation; Theoretical, methodological, and policy issues. In R. W. Marans & D. Stokols (Eds.), *Environmental Simulation; Research and Policy Issues* (pp. 3-21). New York: Plenum Press.

Ulrich, R., Simons, R., Losito, B., Fiorito, E., Miles, M., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology, 11*, 201-230.

Venturino, M., & Wells, M. J. (1990). Head movement as a function of field of view size on a helmet-mounted display. *Proceedings of the Human-Factors and Ergonomics Society, 34th Annual Meeting, 2*, 1572-1576.

Witmer, B. G., Bailey, J. H., Knerr, B. W., & Parsons, K. C. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human-Computer Studies, 45,* 413-428.

Zube, E. H. & Simcox, D. E. (1993). Landscape simulation; Review and potential. In R. W. Marans & D. Stokols (Eds.), *Environmental Simulation; Research and Policy Issues* (pp. 253-278). New York: Plenum Press.