Virtual Experience Test: A Virtual Environment Evaluation Questionnaire

Dustin B. Chertoff *
Intelligent Automation, Inc.

Brian Goldiez †
University of Central Florida

Joseph J. LaViola, Jr. ‡
University of Central Florida

ABSTRACT

We present the development and evaluation of the Virtual Experience Test (VET). The VET is a survey instrument used to measure holistic virtual environment experiences based upon the five dimensions of experiential design: sensory, cognitive, affective, active, and relational. Experiential Design (ED) is a holistic approach to enhance presence in virtual environments that goes beyond existing presence theory (i.e. a focus on the sensory aspects of VE experiences) to include affective and cognitive factors.

To evaluate the VET, 62 participants played the commercial video game *Mirror's Edge*. After gameplay both the VET and the ITC-Sense of Presence Inventory (ITC-SOPI) were administered. A principal component analysis was performed on the VET and it was determined that the actual question clustering coincided with the proposed dimensions of experiential design. Furthermore, scores from the VET were shown to have a significant relationship with presence scores on the ITC-SOPI. The results of this research produced a validated measure of holistic experience that could be used to evaluate virtual environments. Furthermore, our experiment indicates that virtual environments utilizing holistic designs can result in significantly higher presence.

KEYWORDS: Presence, Experiential Design, Virtual Environments

INDEX TERMS: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – Artificial, augmented, virtual realities; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – Evaluation/methodology

1 INTRODUCTION

The experience users have in a virtual space emerges due to a combination of factors. For example, exogenous factors provide the context for a space through the physical and task environment [33]. As a result, the physical environment produces a state of immersion and the task environment can produce a state of involvement. The interaction of these components is associated with the traditional sense of presence, or "being there," in a virtual environment (VE). We use the term VE to represent the combined hardware platform, software systems, and encountered scenarios.

These early models of presence [14][22][37] have largely emphasized the sensory components of experience. Inherent in these models is the belief that a user decision towards the sensory content of a VE indicates that the user became present in that VE. This implies that increasing the sensory fidelity of the VE, be it through the addition of more senses into the experience or through refining a particular sensory channel, is enough to make a person decide to continuously attend to the virtual environment.

However, recent work involving entertainment virtual environments showed that while study participants desired a high-sensory experience, such an experience was "irrelevant 'eye-candy' if the game was not enjoyable" [8]. This notion is enhanced if we think of an experience in the real world. There exist affective and cognitive elements, or endogenous factors, that contribute to a holistic experience beyond the physical location and the tasks performed. There is much emerging support within the literature for a relationship between presence and holistic experiences [7][10][17]. In fact, several researchers have begun to look towards commercial video games to better understand how presence is affected through holistic virtual environments [4][8][19][27]. Researchers have also looked at a sense of flow, or deep involvement, with a virtual environment and immersion [6][13][35].

These recent studies indicate that understanding the factors of overall experience are important for designing better VEs. One promising technique for incorporating overall experience into presence is experiential design (ED) [7]. Experiential design is a design philosophy where various dimensions of experience – sensory, cognitive, affective, active, and relational (see Table 1) – are considered in order to produce a desired user experience [28]. This type of design philosophy has strong roots in the user centered design ideas of [5][26].

In this paper, we introduce a new survey instrument for evaluating the degree a VE incorporates the dimensions of experiential design. This instrument is known as the Virtual Experience Test (VET). The goal of the VET is allow designers to evaluate what aspects of experience their VE needs further work on, such that future iterations can be improved.

In the next section we identify related work on measuring experience in virtual environments. Section 3 reviews the concept of experiential design. Section 4 discusses the development of the VET. Section 5 describes the experiment conducted to evaluate the VET. We present the results of the experiment in section 6. A discussion of the experimental results is in section 7. Future work can be found in section 8. We conclude in section 9.

2 RELATED WORK

The measurement of presence originated with the use of questionnaires regarding the sensory experience of the user. The Slater-Usoh-Steed (SUS) questionnaire consists of six questions that look at three themes to identify a sense of physical presence in an environment [34]. The three themes include: the user's sense of being in the virtual environment, the extent to which the virtual environment becomes the user's primary environment, and the extent to which the virtual environment is remembered as an actual place. The Witmer-Singer Presence Questionnaire (PQ) attempted to look beyond just immersion to measure user involvement as well [37]. The ITC-Sense of Presence Inventory (ITC-SOPI) also looked at the subjective sense of physical presence in addition to other factors related to engagement, ecological validity, and negative effects felt by users in VEs [21].

*e-mail: dchertoff@i-a-i.com
†e-mail: bgoldiez@ist.ucf.edu
‡e-mail: jjl@eecs.ucf.edu

Table 1. Dimensions of experiential design applied towards virtual environments [7]

Experiential Design Dimension	Description
Sensory	Includes sensory input (visual, aural, haptic, etc.) as well as perception of those stimuli. Represented through sensory hardware and software that creates the sensations.
Cognitive	Mental engagement with an experience, such as anticipating outcomes and solving mysteries. Can be interpreted as task engagement.
Affective	Refers to the user's emotional state. Related to the degree to which a person's emotions in the simulated environment would accurately mimic his emotional state in a similar real-world situation.
Active	Relates to the degree of personal connection a person feels to an experience. Associated with the degree of empathy, identification, and personal relation a person feels to the virtual environment's avatars, surroundings, and scenario.
Relational	Comprised of the social aspects of an experience. Operationalized as co-experience; creating and reinforcing meaning through collaborative experiences.

The SUS and PQ were amongst the first widely used presence questionnaires. However, it was noted in [36] by the SUS designers that neither SUS nor PQ could pass a "reality" test. Both the PQ and SUS failed to produce significantly greater presence scores for a person in a real environment than in a virtual one. It was also found by [38] that a lack of reliable statistical validity exists in both the SUS and PQ questionnaires. As a result, researchers have begun measuring causes associated with temporary breaks in presence [32] and whether physiological reactions can be associated with presence [25].

While such measures appear promising, they can be complicated to administer, invasive, and subject to ambiguity. This has led other researchers to evaluate what contributes to experience within high quality commercial video games. This class of VE is consistently labelled highly immersive and involving by users, and could be enlightening on how to design more effective traditional VEs.

Examples of game experience evaluation include the GameFlow heuristics [35] which evaluate player enjoyment in games. GameFlow was subsequently converted into a questionnaire in order to evaluate e-learning game environments [13]. The Game Experience Questionnaire (GEQ) has also been proposed by [19]. The GEQ is based on seven components of game-play experience: immersion, tension, competence, flow, negative affect, positive affect, and challenge.

Further evidence for looking at evaluations of user experience with game environments comes from the Heuristics for Evaluating Playability (HEP) [11]. The HEP was developed around four categories: game play (problems and challenges for the player), game story (plots and character development), game mechanics (rules), and game usability (interface and interaction methods). The authors found that many of the issues found during in-depth user evaluation were also uncovered by the HEP [11].

With issues surrounding the validity of two existing major presence questionnaires, it appears constructive to further investigate alternative presence evaluation techniques.

3 Experiential Design

It was suggested that virtual environments be designed with a user's holistic experience in mind [7]. The aim is to integrate the various elements of experience – sensory, cognitive, affective, active (personal), and relational (social) (see Table 1) – in order to elicit an enhanced sense of presence and to create situations where accurate, memorable, and stable schema can be developed. The process of using these dimensions to create such an experience is known as experiential design (ED) [28].

ED originated from the marketing field where it was used to encourage people to create meaningful emotional and social connections to a product. For example, a person might construct a personal narrative about an experience in a coffee shop listening to live music on a couch. While the product being sold is coffee, the use of music and a homey environment creates a more compelling environment. The person will subsequently create episodic memories about the experience, and if they enjoyed the artifacts of that experience, will build positive associations with the product being sold [3].

Because of its reliance on positive associations stored in memory, a cornerstone of ED is considering how prior, similar experiences can be integrated into new user experiences. This idea is based on the self-referencing effect of cognition [23]. As we encounter new information, we attempt to relate it to experiences stored in memory. Each query triggers a variety of associated memories to be remembered. Thus, a strong memory of an experience can be recalled at a later time when the correct trigger is provided. For example, if our customer hears the same song she might be reminded of the positive experience she had drinking coffee at our shop.

VEs can also make use of this aspect of cognition. When a VE is designed to trigger a corresponding user experience, a variety of missing information can fill in the gaps left by the VE. This has the potential side-effect of increasing the user's sense of presence [31]. By using holistic designs, the opportunity to both trigger a response from and to store experience in a potentially larger network of schema. This could increase the potential for performance increases, a connection sought after by previous presence researchers [2].

4 Virtual Experience Test

Development of the Virtual Experience Test began by creating a series of heuristics based upon the experiential design dimension descriptions. From the sensory dimension, we determined that the heuristics should focus on the consistency and quality of sensory information. From the cognitive dimension, heuristics that assessed the ability to complete a task and to understand the environment's rules were created. Heuristics from the affective dimension measured the strength and variety of emotion towards the environment. The active dimension heuristics consisted of the ability to become the avatar in the environment, the use of story to explain the environment's content, and the ability to reuse skills throughout the environment. Heuristics from the relational dimension would measure the quality of interactions with agents and other users. After the initial heuristic categorization, further refinements were made to capture previously identified associations with presence factors. These associations are now explained further.

The **sensory dimension** consisted of seven heuristics regarding how well the senses were utilized during interactions with the VE. A review done by [20] showed that both the quality of sensory hardware and the sensory content had a widespread positive effect on reported presence. Thus, heuristics were included to address sensory hardware and content quality. In addition, the consistency of sensory information is rated.

The **cognitive dimension** produced five heuristics. These heuristics focused on how well the environment supported task engagement through the clarity of task explanations, perceived task interest, explanation of environment rules, and the ability of the environment to support multiple solutions for a task. The basis for these heuristics came from the identification of various environmental control factors shown to affect presence [18][30][37] and their similarity to the flow experience [9].

The **affective dimension** contained four heuristics. These heuristics focused on the expected emotional impact of the user towards the included scenarios. This included the variety, strength, and relevance of emotions experienced while completing the scenarios. This set of heuristics also looked at how the environment conveyed desired user emotions through dialog, non-verbal cues (i.e. through agent postures and facial expressions), and audio. A relationship between presence and a high sense of arousal [14][15][25], as well as between high user enjoyment and higher reported presence [1] have been shown, lending support for the impact of the affective dimension on presence.

The **active dimension** also consisted of four heuristics. These heuristics were concerned with the expected level of user attachment towards the VE. Attachment was defined as the degree to which the user would think they were a character in the environment, the level of content reuse, and the utilization of narrative. A compelling argument for presence being fundamentally tied to agency and environment control has been made [16]. Further, there is support that narrative based virtual environments have an impact on increased presence [24].

The **relational dimension** contained four heuristics. These heuristics were concerned with the social aspects of the environment. They focused on the expected level and quality of user interactions with agents in the environment. Previous research had shown that presence in social situations was higher, be it with agents, strangers, or friends [14][29].

Based upon these identified associations, it was hypothesized that evaluating a VE using the heuristics would allow for predictions to be made regarding how presence inducing the environment would be. An iterative process was then employed to convert the list of heuristics into a questionnaire. Each question was rated on a 5-point Likert scale, with a 1 indicating "strongly disagree" and a 5 indicating "strongly agree." Two participants familiar with gaming and virtual environments, but not the theory surrounding presence and experiential design, were asked to pilot the questionnaire. These participants were asked to explain what they thought each question was asking and to identify any unfamiliar terminology, or that they thought would be unfamiliar to a novice. Suggestions by the participant were noted and used in the next iteration of the questionnaire. The resulting questions can be found in the first column of Table 2. The second column contains the question's associated dimension of experiential design. The third column contains the grouping established after questionnaire analysis was completed. The details of this column are explained further in section 6.

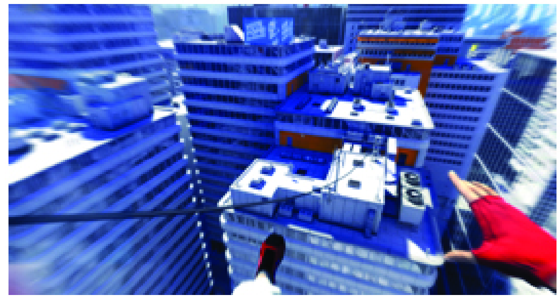

Figure 1. Exploring the world of *Mirror's Edge*

5 VET Evaluation

Although the VET was heavily based on existing presence and experiential design theory, it still required validation to ensure that the chosen questions adequately related to their proposed dimensions. A study was therefore designed to investigate the validity of the VET. If the proposed questions and dimensions were in fact valid, then it would also be necessary to determine if the VET could predict presence in different VEs. To ensure the VET was capable of doing so, an existing VE that incorporated the dimensions of experiential design in different configurations was needed. To this end, the commercial game *Mirror's Edge*™ was selected, as it contained two game configurations that utilized the dimensions of experiential design in different ways. If the VET was capable of predicting presence, then differences on the VET scores for each game configuration would also be reflected in differences on presence scores. This question formed the basis of the study explained in the rest of this section.

5.1 Methodology

In *Mirror's Edge* the player takes control of a character that runs, jumps, and climbs around various rooftops and interiors from a first-person viewpoint [12]. The character's arms and feet were visible when the player moved around (see Figure 1). With its focus on exploring and interacting, *Mirror's Edge* provided a similar experience to traditional egocentric, walkthrough style virtual reality systems. In addition, there was a strong story component to the world, which made it a good hybrid between traditional virtual reality and gaming.

Mirror's Edge included two game-types. The first game-type was a story-mode. This game mode included a narrative and additional characters that the player interacted with. The second mode was a time-trial mode, where the player travelled across a map to interact with various way-points in a limited amount of time. Based on an analysis by the authors, it was determined that the two game types sufficiently differed in how the dimensions of experiential design were utilized regarding the affective, active, and relational dimensions.

This led to two hypotheses.

Hypothesis 1: The story-mode game-type of *Mirror's Edge* would receive higher experiential design scores than the time-trial game-type in the affective, active, and relational dimensions.

As a result of the story-mode game-type providing more elements of experience, the next hypothesis was produced.

Hypothesis 2: The story-mode game-type of *Mirror's Edge* would receive higher presence scores.

Table 2. Virtual Experience Tool questions with the associated original experiential design dimension and the factor determined after principal component analysis (see section 6.1)

	Question	Experiential Design Dimension	Post-PCA Factor
1)	**I found the visual display hardware to be of high quality.**	Sensory	-
2)	**I experienced a high level of interaction with computer agents in the virtual environment.**	Relational	Story Telling
3)	**I found the visual content of the environment to be of high quality.**	Sensory	-
4)	**I think that the environment was able to support multiple human users at the same time.**	Relational	-
5)	**When I felt an emotional reaction, I felt that my emotional state was appropriate given the events that occurred in the virtual environment at that time.**	Affective	Story Telling
6)	**I found that the virtual environment did a good job of using a story to explain my tasks.**	Active	Story Telling
7)	**I felt a variety of emotions while working on the environment's tasks.**	Affective	Story Telling
8)	**I found that a high level of interaction with other users or computer agents was required in order to complete my tasks in the virtual environment.**	Relational	Story Telling
9)	**I felt that computer controlled (artificial intelligence) agents were used well in the virtual environment.**	Relational	Story Telling
10)	**I had an emotional reaction while working on the environment's tasks.**	Affective	-
11)	**I believed that I was the character I was controlling.**	Active	Active
12)	**I found that the content in the virtual environment was helpful in informing me of my current task.**	Cognitive	Task Completion
13)	**I feel that I could construct a story about my actions in the environment.**	Active	Active
14)	**I found the user interface to be helpful in informing me of my current task.**	Cognitive	Task Completion
15)	**I found the haptic content of the environment to be of high quality (haptics refers to the sense of touch).**	Sensory	Haptics
16)	**I thought that the virtual environment made it clear what I was and was not allowed to do.**	Cognitive	-
17)	**I found the audio hardware to be of high quality.**	Sensory	-
18)	**I felt that the environment used multiple techniques to convey emotion.**	Affective	Story Telling
19)	**I found the audio content of the environment to be of high quality.**	Sensory	Sensory Content
20)	**I thought that the tasks I was able to do in the virtual environment were interesting.**	Cognitive	Sensory Content
21)	**I felt that the virtual environment allowed me to complete my task in several different ways.**	Cognitive	-
22)	**I felt that I was able to continuously reuse techniques that I learned on previous tasks on my later tasks.**	Active	Active
23)	**I found the haptic hardware to be of high quality (haptics refers to the sense of touch).**	Sensory	Haptics
24)	**I found that the sensory information of the virtual environment was consistent. For example, the sound of two metal objects colliding sounded metallic. A visually smooth object felt smooth.**	Sensory	Sensory Content
	A dash (-) denotes the question was removed based upon the results of the principal component analysis process		

5.2 Participants and Procedure

A total of 62 participants (52 male, 10 female) were recruited from the general campus population. Mean participant age was 23 with a standard deviation of 4.13. Participants were primarily recruited from the honors, engineering, and digital media colleges of the University of Central Florida; however no restrictions on background were imposed. Participants were awarded $10 after the completion of the study. Participants were provided an informed consent form discussing the possible effects of participation in the study. Additionally, participants were informed that they could stop at any time during the experiment.

Various participant demographics were collected including computer and game experience. All but 3 participants reported intermediate or higher experience with computers. 40 of the 62 participants reported playing games often (multiple times a week) or more. The majority of participants also had a basic understanding of how computer/video game graphics were produced and basic knowledge about virtual reality.

Participants first played an in-game tutorial to learn the basic rules and controls of the game. During the tutorial, the experimenter answered questions about game mechanics and provided advice on how to perform more advanced techniques. After the tutorial, participants were told to complete several tasks in at most 45 minutes in either the story-mode or time-trial game-types. The game-type played was randomly assigned. After playing the game, participants filled out the ITC-SOPI [21], the VET, and a modified version of the Experience Sampling Method (ESM), which was used to measure a state of flow [9]. The results of the ESM analysis are not, however, reported in this paper.

Questionnaire order was randomized using a 3x3 Latin square to reduce questionnaire order bias. After the final questionnaire was finished, participants were thanked and given $10 for their participation. Any questions the participant had about the study were then answered.

5.3 Apparatus and Test Environment

All game-play was performed on the Microsoft XBOX 360 Elite gaming console using a wireless controller. The console was connected to a 42" EDTV 480p (852 x 480 resolution) plasma TV. Participants were seated about 6 feet from the screen. The room was kept lit to be consistent with a typical gaming experience. The door to the room was kept closed during the study to minimize extraneous sound.

In the story-mode of *Mirror's Edge*, the participant was first tasked with navigating the environment to find a character. Navigation of the environment included jumping over obstacles and between buildings, along with climbing up walls (see Figure 1). During several points, the participant had to choose between fighting and avoiding enemies that were trying to shoot the participant's character. The participant was then tasked with escaping a building while being pursued by enemy characters. The escape took the character through a variety of locations such as the inside of a building, several building roofs, and a subway station.

In the time-trial mode, the participant was tasked with racing from one map waypoint to another in under a certain amount of time. There were no characters other than the participant's avatar in the time-trial environments. Two time-trial maps were used. The first map was identical to the in-game tutorial. Participants had to reach all of the way-points in less than 2 minutes. If they did not, they were told to restart the map and try again until they could successfully navigate in the allotted time. Once successful, the participant would move on to the second map. The second map used the same environment as the first map, but way-points were in a different order. The participant had 2 minutes and 10 seconds to complete this map. Again, participants were told to restart the map if they were unable to complete it in the allotted time.

Both of these game-types were expected to be high-immersion, high-involvement. However, different design techniques were used to produce such feelings. For either condition, participants were told to stop playing at the 45 minute mark if they had not yet completed all of the game tasks assigned.

6 Results

6.1 VET Analysis

The VET consisted of 24 questions based upon the five dimensions of experiential design: sensory, cognitive, affective, active, and relational. Thus, there was an expectation that certain questions would be correlated and would be grouped by a subsequent principal component analysis. Therefore, the first step in verifying the design of the VET was to determine if groupings based on a statistical analysis matched the predicted groupings from experiential design.

We analyzed the VET by first visually inspecting the correlation matrix of the original 24 questions. Questions that only correlated to 4 or less other questions were removed. This resulted in 6 questions being eliminated. Next, questions that were highly correlated were analyzed. Only one pair of questions (questions 7 and 10) was found to be highly correlated. Both Q7 and Q10 asked about whether the participant felt an emotional reaction. Q10 was eliminated as it correlated to fewer other questions than Q7. Ultimately, 7 questions were eliminated, resulting in 17 questions being used during a principal component analysis (see Table 2).

KMO and Bartlett's Tests were performed to ensure that the remaining data was sufficient to proceed with the principal component analysis (PCA). The KMO test yielded a value of 0.75. The Bartlett's Test of Sphericity was not violated ($p < 0.05$). In addition, the determinate of the correlation matrix was 0.001. These values all suggest that the PCA could be successfully performed.

Factors with Eigenvalues greater than 1.0 were extracted by the PCA. An inspection of the resulting Scree plot confirmed the extraction of only 5 factors. The resulting factors were then rotated using a Varimax rotation (a type of orthogonal rotation). The five rotated factors accounted for 65.8% of the variance.

Factor 1 (Story-Telling) consisted of questions 2, 5, 6, 7, 8, and 18. Questions 2 and 8 were originally part of the relational dimension. Questions 5, 7, and 18 were from the affective dimension, and question 6 was from the active dimension. These questions were all related to the communication of and interaction with the narrative elements of a virtual environment.

Factor 2 (Haptics) consisted of questions 9, 15, and 23. Question 9 was expected to be part of the relational dimension. Questions 15 and 23 both dealt with haptics and are part of the sensory dimension. While Q9 loaded highest on the Haptics factor, its content was most relevant to the Story-Telling factor. The correlations between Q9 and the other questions in the Story-Telling factor supported this. In addition, prior to rotating the component matrix, Q9 loaded highest on the Story-Telling factor. Q9 was therefore moved to the Story-Telling factor. As a result, factor 2 was related to the utilization of haptics in the VE.

Factor 3 (Sensory Content) included questions 19, 20, and 24. Questions 19 and 24 were originally part of the sensory dimension, while question 20 was from the cognitive dimension. While it appears out of place, Q20 does relate to the other questions in the grouping, as sensory elements are involved in both communicating and completing tasks. This indicates factor 3 is concerned with non-haptic sensory content in the environment.

Factor 4 (Task Completion) included questions 12 and 14. Both of these questions came from the cognitive dimension and related to the level of help participants received on their tasks from both the environment and the user interface. Factor 4 is therefore concerned with task completion.

Factor 5 (Active) included questions 11, 13, and 22. All of these questions were part of the active dimension of experience. Thus, factor 5 is concerned with the degree to which the participant felt that they were the character in the environment.

The question text associated with each of these five factors can be found in Table 2.

6.2 Game Condition and VET Scores

Recall that Hypothesis 1 stated that the story condition of the game would receive higher experiential design scores than the time-trial condition. Taking into account the new factors produced, it was expected that the story-mode condition would see high scores in the story telling and active factors. This was because the story telling factor constituted questions that were originally in the affective and relational dimensions of experiential design, and the active factor contains questions from the active dimension of experiential design. No claims regarding significant differences in the other factors were made.

Table 3. VET factor estimated means by condition. Means range from 1 (strongly disagree) to 5 (strongly agree).

VET Factor	Condition	Estimated Mean	Std. Error
Story Telling	Story	3.595	.137
	Time	3.019	.137
Haptics	Story	3.523	.170
	Time	3.203	.170
Sensory Content	Story	4.371	.105
	Time	3.888	.105
Task Completion	Story	3.890	.137
	Time	3.561	.137
Active	Story	3.925	.125
	Time	3.473	.125

Table 4. Presence factor estimated means by condition. Means range from 1 (strongly disagree) to 5 (strongly agree).

Presence Factor	Condition	Estimated Mean	Std. Error
Spatial Presence	Story	3.120	.107
	Time	2.919	.107
Engagement	Story	3.981	.099
	Time	3.632	.099
Ecological Validity/ Naturalness	Story	3.028	.145
	Time	3.108	.145
Negative Effects	Story	2.019	.147
	Time	1.811	.147

As the VET measures holistic experience, each experience factor was expected to be related to the others. This was confirmed by looking at the resulting correlation matrix for VET scores on each factor. This suggested that a multivariate analysis was needed. Further, previous game experience could have potentially affected the results, so the previous game experience question of the ITC-SOPI demographics was used as a covariate. Therefore, a single MANCOVA could be performed.

VET questions were scored from 1 (strongly disagree) to 5 (strongly agree). Estimated means were calculated for each VET factor using a game experience mean of 2.13 (see Table 3). Previous game experience was not found to have a significant effect. Statistically significant differences were found for the story telling ($F(1,59) = 8.698$, $p < 0.01$), sensory content ($F(1,59) = 10.38$, $p < 0.01$) and active ($F(1,59) = 6.455$, $p < 0.05$) factors. No significant differences in the haptics and task completion factors were observed. Note that the covariate accounted for an additional degree of freedom (resulting in a total of 61 degrees of freedom between covariate, treatment, and error).

Hypothesis 1 was supported with significant differences in the story telling and active factors, with no difference in the haptics and task completion factors. However, there was a significant difference in the sensory content factor. The observed sensory content difference suggests that the different sensory content of the story-mode did have an effect on experience. Essentially, the more holistic design of the story-mode condition resulted in a better experience than the time-trial condition. These findings support the notion of using experiential design techniques.

6.3 Effect of Condition on Presence

Presence was measured using the ITC-SOPI, which produces four scores related to presence: spatial presence score (SPS), engagement score, ecological validity/naturalness score (EVNS), and negative effects score (NES). Recall that in Hypothesis 2, it was expected that the story-mode condition would receive higher presence scores than the time-trial condition.

All but the NES were found to be highly correlated to the other scores. As a result, a single MANCOVA was again conducted comparing condition and underlying ITC-SOPI presence scores. Estimated means for each presence score are presented in Table 4. Previous game experience did not have a significant effect on ITC-SOPI scores. A statistically significant difference was found for the engagement score ($F(1,59) = 6.054$, $p < 0.05$). No significant differences were observed for the SPS, ENVS, or NES.

Given the addition of an interactive narrative element in the story-mode, the increase in engagement is not surprising. This suggests participants had a better experience in the story-mode condition. Further support for this claim comes from the significantly higher experience scores observed in the story-mode condition reported in section 6.2. No time-trial mode scores were observed to be significantly greater than the corresponding story-mode scores. With the increase in scores favoring the story-mode condition, Hypothesis 2 can be partially accepted.

6.4 VET – Presence Relationship

The findings regarding Hypotheses 1 and 2 indicate that some type of relationship exists between holistic designs and the participant's experience. To further explore this relationship, we looked at whether scores from the VET significantly related to presence scores. An overall VET score was calculated by taking the average of the five VET factors, which smoothed out potential outlier scores. This results in a score ranging from 1 to 5. The overall score was then recoded such that a score between 1 and 2 (not inclusive) was "Very Low Experience," between 2 and 3 (not inclusive) "Low Experience," between 3 and 4 (not inclusive) "High Experience" and greater than 4 "Very High Experience." With this categorization, we could analyze whether high scores on the VET corresponded to higher presence factor scores.

A single MANCOVA was performed comparing presence factor scores with overall VET scores (see Table 5 for estimated means). A statistically significant effect was found between the VET and presence for the spatial presence score ($F(2,58) = 3.768$, $p < 0.01$), engagement score ($F(2,58) = 4.535$, $p < 0.01$), and ecological validity/naturalness score ($F(2,58) = 5.325$, $p < 0.01$). No significant difference was found for the negative effects score. These results indicate that the VET is capable of predicting increases in presence as measured by the ITC-SOPI and support the use of the VET to predict potential presence in future studies.

7 Discussion

The VET originally consisted of 24 questions based on the dimensions of experiential design. Existing theory and empirical results regarding presence were also used in the question generation process. Essentially, prior to the principal component analysis of the VET being performed, a general idea of how the questions should group was known. The factors produced after the factor analysis suggests that much of the initial idea was correct, as both a cognitive dimension and active dimension were shown to exist. However, there were two exceptions.

First, the factor analysis combined the affective and relational dimensions into a single factor. It is possible that the relational dimension can be seen as the communication medium for the

Table 5. Estimated presence factor means for participants with the "low," "high," and "very high" VET scores.

Presence Factor	VET Category	Estimated Mean	Std. Error
Spatial Presence	Low Exp	2.363	.185
	High Exp	2.965	.074
	Very High Exp	3.654	.148
Engagement	Low Exp	2.826	.162
	High Exp	3.848	.064
	Very High Exp	4.264	.129
Ecological Validity/ Naturalness	Low Exp	2.417	.261
	High Exp	2.968	.104
	Very High Exp	3.882	.208
Negative Effects	Low Exp	1.821	.309
	High Exp	1.997	.123
	Very High Exp	1.645	.246

affective components of the environment. Story-telling typically involves the description of character actions to produce an emotional response in the receiver of that story. Therefore the inclusion of questions regarding the actions of artificial intelligence characters on the same factor with questions about the emotional impact of those actions is consistent. In other words, this suggests that the affective and relational dimensions of experience are very strongly coupled and could potentially be viewed under a single lens of interactive narrative.

It is also possible that this combination is a product of the virtual environment used. All computer controlled agents acted as communicators of narrative or as elements of the narrative. Thus, emotional reactions to the story or tasks involved the very elements the relational dimension intended to explore. In other words, emotional reactions were the product of interacting with computer agents. This is reinforced by the significant difference between the story-mode and time-trial conditions for the story-telling factor. Without the presence of agents in the time-trial, the only emotional reaction felt was due to success or failure with the waypoint navigation task.

Second, separate sensory factors emerged. It is interesting that haptics had their own dimension, but that is very likely due to the fact that very little haptic interaction was a part of the game. The controller would vibrate when the participant's character was shot, but not for any other physical interactions in the game, such as landing after a fall or grabbing onto a building. It is possible that had the game made better use of the rumble features of the controller, the questions in factors 2 and 3 would have been found to belong to a single factor. It is suggested that those doing future studies involving the VET tentatively keep this in mind as they explore how coupled the two factors are.

There is one other aspect of the VET that should be noted. The questions regarding participant emotion do not distinguish between internally and externally derived emotions. For example, feelings of boredom and frustration are emotions derived from an internal perception of performance on a task. These emotions are different from those due to sympathy/empathy to the plights of characters. As a result, it might prove difficult to pull apart what is an emotional reaction to an environment's narrative, and what is a reaction to task performance. If the goal is to simply tell if any type of emotional reaction occurred, then no issue is foreseen. However, if one wants to determine if the emotional reaction is due to internal or external factors, additional steps would be required.

The comparison of VET scores based on condition produced an unexpected result in that the sensory content score of the story-mode was significantly higher than in the time-trial mode. The environments in each condition used a visually consistent style and the same sound effects, so no significant differences were expected. However, the observed difference in sensory content scores indicates that the manner in which the sensory content was used in the story-mode condition had a positive effect on experience. This could be due to several things, such as more varied environments (inside buildings, outside buildings, on rooftops), cut-scenes, mood music, and voice communication between characters over a radio. This indicates that the story-mode made a more concerted effort to incorporate the sensory content to the other dimensions of experience, producing a more holistic design. These results are consistent with the work of [8] who found that the level of sensory immersion depended on how incorporated the sensory elements were on goal completion. Taken together, there is support for the future use of experiential design techniques for virtual environments.

8 Future Work

As only one environment was studied, the reliability of the VET across multiple environment types remains to be determined. While the VET is sensitive to the differences in *Mirror's Edge* game-modes, evaluations of other environments are still needed. In particular, questionnaire validation should be performed with other environments to ensure principal component analysis consistency. As noted, the distinction between the Haptics and Sensory Content factors are believed to be an artifact of the game used in this study. Thus, a study using a virtual environment with more tightly integrated haptic feedback should be conducted.

Further, VET reliability remains to be determined. While the questions making up the VET emerged from a variety of virtual environment studies, a future study is needed to ensure that the VET is reliable across a multitude of environments and hardware configurations. One possibility is to choose two games from the same genre to see if the VET can reliably rate that class of environment. Alternatively, the same game could be compared on different hardware platforms.

Test-retest reliability also remains to be determined. This would be accomplished by recruiting the same participants to evaluate their experience on a second occasion.

As the results of this study indicate that holistic environments have the potential for increasing presence, an additional avenue of research would be to analyze the ability of a participant to stay focused in a holistic environment when various stimuli are competing for attention. It is expected that it would be harder to break away the participant enjoying a higher quality experience. However, the stimuli strength required to degrade the participant's experience is unknown. The results of such a study could then fuel guidelines for what types of designs are most relevant for an expected level of possibly competing stimuli.

9 Conclusion

The results of this study illustrate how holistic designs of virtual environments contribute to increased sense of presence. Based on the work of this study and others before, it can be clearly seen that non-sensory components of experience are also related to presence. As more of the population becomes exposed to types of virtual environments, especially games, a certain overall quality of experience will become expected. In order to create VEs that meet these expectations, tools that assist researchers in the design process are needed. Based on both presence and experience theory, the VET will allow researchers to continuously evaluate their world through a variety of experience dimensions and can iterate their aspects of their designs until a holistic experience is achieved.

As one of the goals of presence research is to investigate the concept of "being there," it is important that the VEs studied offer the same holistic experiences attainable in the real world. It is believed that an evaluation tool such as the VET will help meet this challenge.

ACKNOWLEDGEMENTS

This work is supported in part by NSF CAREER award IIS-0845921 and NSF Award IIS-0856045. We wish to thank the anonymous reviewers for their valuable suggestions.

REFERENCES

[1] Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. In G. Salvendy, & M. J. Smith, *Human-computer interaction: Software and hardware interfaces* (pp. 669-704). Amsterdam: Elsevier.

[2] Barfield, W., Zeltzer, D., Sheridan, T., & Slater, M. (1995). Pressence and performance within virtual environments. In W. Barfield, & T. A. Furness, *Virtual environments and advanced interface design.* New York, NY, USA: Oxford University Press.

[3] Battarbee, K., & Mattelmaki, T. (2002). Meaningful product relationships. *Proceedings of the 3rd Conference of Design and Emotion.* Loughborough, England.

[4] Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. *Proceedings of CHI 2004*, (pp. 1297-1300). Vienna, Austria.

[5] Buxton, B. (2007). Sketching User Experiences. San Francisco, CA: Morgan Kaufmann

[6] Chen, X. (2006). *Flow in Games.* University of South California.

[7] Chertoff, D. B., Schatz, S. L., McDaniel, R., & Bowers, C. A. (2008). Improving presence theory through experiential design. *Presence: Teleoperators and Virtual Environments , 17* (4), 405-413.

[8] Clarke, D., & Duimering, P. R. (2006). How computer gamers experience the game situation: A behavioral study. *ACM Computers in Entertainment , 4* (3), 1-23.

[9] Czikszentmihalyi, M., & Czikszentmihalyi, I. S. (1992). *Optimal experience: psychological studies of flow in consciousness*. New York: Cambridge University Press.

[10] Davide, F., & Walker, R. (2003). Engineering Presence: an Experimental Strategy. In G. Riva, F. Davide, & W. A. IJSselsteijn, *Being There: Concepts, effects and measurement of user presence in synthetic environments* (pp. 41-57). Amsterdam, The Netherlands: Ios Press.

[11] Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. *Proceedings of CHI 2004*, (pp. 1509-1512). Vienna, Austria.

[12] Electronic Arts. (2008). *Mirror's Edge*. http://www.mirrorsedge.com/.

[13] Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). EGameFlow: A scale to measure learner's enjoyment of e-learning games. *Computers and Education , 52*, 101-112.

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

[15] Heeter, C. (1995). Communication research on consumer VR. In F. Biocca, & M. R. Levy, *Communication in the age of virtual reality* (pp. 191-218). Mahwah, NJ: Erlbaum.

[16] Herrera, G., Jordan, R., & Vera, L. (2006). Agency and presence: A common dependence on subjectivity? *Presence: Teleoperators and Virtual Environments , 15* (5), 539-552.

[17] IJsselsteijn, W. A. (2003). Presence in the Past: what can we learn from Media History? In G. Riva, F. Davide, & W. A. Ijsselsteijn, *Being There: Concepts, effects and measurement of user presence in synthetic environments* (pp. 17-40). Amsterdam, The Netherlands: Ios Press.

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

[19] IJsselsteijn, W., de Kort, Y., Poels, K., Jurgelionis, A., & Bellotti, F. (2007). Characterising and measuring user experiences in digital games. *ACE Conference 2007.* Salzburg, Austria.

[20] Lee, K. M. (2004). Why presence occurs: evolutionary psychology, media equation, and presence. *Presence: Teleoperators and Virtual Environments , 13*, 494-505.

[21] Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments , 10*, 282-298.

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

[23] Matlin, M. W. (1998). *Cognition* (4th ed.). Fort Worth, Texas, USA: Harcourt Brace College Publishers.

[24] McQuiggan, S. W., Rowe, J. P., & Lester, J. C. (2008). The effects of empathetic virtual characters on presence in narrative-centered learning environments. *Proceedings of the 26th SIGCHI Conference*, (pp. 1511-1520). Florence, Italy.

[25] Meehan, M., Insko, B., Whitton, M., & Brooks, F. P. (2001). Physiological measures of presence. *Proceedings of the 4th International Workshop on Presence.* Philadelphia, Pennsylvania, USA.

[26] Norman, D. (1990). *The Design of Everyday Things.* New York, NY, USA: Doubleday.

[27] Pinchbeck, D. M. (2005). Is presence a relevant or useful construct in designing game environments? *Proceedings of the 3rd International Conference in Computer Game Design and Technology.* Liverpool, UK.

[28] Pine, J. & Gilmore, J. (1999). *The Experience Economy.* Boston, MA, USA: Harvard Business School Press.

[29] Ravaja, N., Saari, T., Turpeinen, M., Laarni, J., Salminen, M., & Kivikangas, M. (2006). Spatial presence and emotions during video game playing: does it matter with whom you play? *Presence , 15* (4), 381-392.

[30] Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments , 1*, 120-125.

[31] Slater, M. (2002). Presence and the sixth sense. *Presence: Teleoperators and Virtual Environments , 11*, 435-439.

[32] Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., et al. (2006). Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments , 15*, 553-569.

[33] Slater, M., & Usoh, M. (1993). Presence in Immersive Virtual Environments. *IEEE Virtual Reality Annual International Symposium*, (pp. 90-96).

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

[35] Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *Computers in Entertainment , 3* (3).

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

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

[38] Youngblut, C., & Perrin, B. M. (2002). Investigating the relationship between presence and performance in virtual environments. *IMAGE 2002 Conference.* Arizona, USA.