Towards an Explanatory Design Theory for Context-dependent Learning in Immersive Virtual Reality

Research-in-Progress

Katarina Jahn, Henrik Kampling, Hans Christian Klein, Yasin Kuru, Bjoern Niehaves

Abstract

Immersive virtual reality (IVR) is increasingly used for learning. However, research on specific designs for IVRs which can be used to enhance individual learning performance is still at an early stage. In this research-in-progress paper, we build upon theories on context-dependent learning to develop an explanatory design theory. We hypothesize that if the user learns in a virtual environment that represents the recall environment (environmental congruence), recall is facilitated. Additionally, if the IVR is designed with a high degree of sensory immersion, the effect of environmental congruence on learning is further increased through enhanced cognitive absorption in the technology. In contrast, cognitive absorption in the task should have a reversed effect. To test the explanatory design theory, we plan to conduct a 2 (learning environment: Room A vs. Room B) x 2 (sensory immersion: low vs. high) x 2 (recall environment: Room A vs. Room B) between-subjects laboratory experiment.

Keywords: immersive virtual reality, cognitive absorption, context-dependent learning, place-dependent learning, explanatory design theory, design science, laboratory experiment

Introduction

Forms of immersive virtual reality (IVR), a technology in which the user is completely absorbed into by the use of head-mounted displays, are increasingly used for learning in different contexts. There are IVR applications used for learning in schools, universities and in health care (Martín-Gutiérrez et al. 2017). Additionally, organizations such as VW started to use IVR for letting their employees learn new organizational processes (Hayden 2018).

IVR has not only the advantage that learning can be designed highly engaging by involving the learner deeply into what they are doing, but also that it can be used to re-create places that are not available to the learner. For learning, the latter can be especially beneficial because of environmental context-dependent memory. According to research on environmental context-dependent memories (Isarida and Isarida, 2014), learning and recalling in the same place is more beneficial for individual learning performance than learning and recalling in different places. However, to the best of our knowledge, there is no research on how environmental context-dependent memory effects can be recreated. With the research-in-progress paper at hand, we want to address the following research question.

RQ: How can IVR be designed to enhance context-dependent learning when it is not possible to learn in the environment where recall takes place?

Explanatory design theories (Baskerville and Pries-Heje 2010; Gregor 2009; Kuechler and Vaishnavi 2012; Niehaves and Ortbach 2016) can answer this research question by not only stating how to design an artifact, but also explain why specific design options have specific effects (Gregor 2009; Kuechler and Vaishnavi 2012) through the use of structural equation modeling terminology (Niehaves and Ortbach 2016). In this research-in-progress paper, we draw upon theories on environmental context-dependent memory and cognitive absorption (CA) to develop an explanatory design theory for context-dependent learning that answers our research question (see Figure 1). We propose that environmental congruence enhances individual learning performance and that this effect is further increased by a high degree of sensory immersion which increases CA in the technology through heightened presence. We plan to conduct a 2 (learning environment: Room A vs. Room B) x 2 (sensory immersion: low vs. high) x 2 (recall environment: Room A vs. Room B) between-subjects experiment to test the explanatory design theory.

Theoretical Background and Model Development

Environmental Context-Dependent Learning

Theories about context-dependent memory (Isarida and Isarida 2014; Smith and Vela 2001) state that different contextual cues can affect recall of target information. Whereas target information is defined as the information that should be remembered, contextual cues represent information that is not the target information but was present (physically or mentally) during encoding. If the context is encoded with the target information, the context can be used as retrieval cue for remembering (Isarida and Isarida 2014). Therefore, the learning performance of individuals (individual learning performance, ILP) can be enhanced through the use of context-dependent learning.

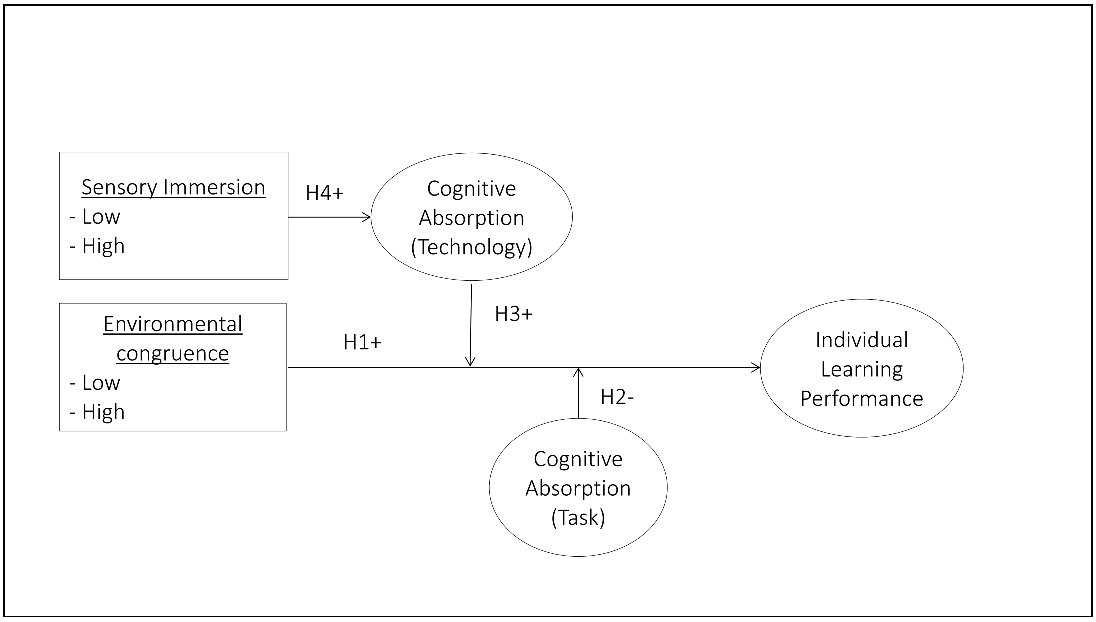

Figure 1. Explanatory Design Theory for Context-dependent Learning

Theories about *environmental* context-dependent memory specify this effect for aspects related to the environment in which the target information was learned (Isarida and Isarida 2014). Environmental context can consist of the larger environment, such as place (Smith and Vela 2001) or specific aspects of the environment, such as odor (Isarida et al. 2014), background color (Isarida and Isarin 2007), or background music (Isarida et al. 2017). For example, when an employee learns how to use a machine, new to the organization in an office, different elements of the learning environment (e.g. lightning, desktop computer, background music) can be encoded with the target information during learning. Therefore, recall of the target information might be hindered, when the employee tries to remember it in a different learning environment, such as a production hall, which consists of different environmental elements.

With physical reinstatement of the environment in which the learning took place, recall is facilitated (Godden and Baddeley 1975; Isarida and Isarida 2014; Smith et al. 1978). In the case of the employee, if they return to the office, the physical reinstatement of the environmental context could enhance recall, because the elements of the environmental context were encoded together with the target information. However, physical reinstatement is often difficult in practice. Returning to the office for recall every time the machine gives an error message which the employee learned at the office would be time-consuming. Additionally, returning to the office might not be a helpful solution if the employee has to remember the information in the production hall while using the machine and not in the office. Therefore, an alternative to physical reinstatement would be mental reinstatement – the mental visualization of the environmental learning context. However, mental reinstatement can be too difficult in some circumstances (Canas and Nelson 1986) or needs to be requested explicitly in order to be used for some individuals (e.g. older adults, Fernández and Alonso 2001).

With the use of IVR, it is possible to simulate the context almost completely without being physically located in the initial learning environment. Thus, with the use of IVR in the learning situation, employees could benefit from the advantages of physical reinstatement without the costs of mental reinstatement. However, to the best of our knowledge, research on using IVR in the learning situation has not yet been done. Based on the described place-dependent memory effects in real learning and recall environments, we assume that learning and recall in a congruent environment is more beneficial to the user than learning und recall in an incongruent environment. Regarding the design, we therefore hypothesize that learning in a room that is congruent to the recall room is more beneficial for recall than learning in a room that is incongruent to the recall room.

H1: Environmental congruence leads to a higher ILP than environmental incongruence.

Cognitive Absorption

In the field of learning, cognitive absorption (CA), also called flow, refers to a state in which an individual is completely involved with a task (CA_TASK, Csikszentmihalyi 1990). CA_TASK is usually identified to be a desirable state for learning (see Table 1 for an overview of construct definitions). However, some research on CA_TASK suggested that high levels can be detrimental because context effects are blocked out (Magni et al. 2013). Likewise, research on environmental context-dependent learning has indicated that a high involvement with the task, and therefore high CA_TASK, decreases ILP because the environment is blocked out (Smith and Vela 2001). We therefore hypothesize that CA_TASK moderates the effect of environmental congruence on ILP.

H2: CA_TASK moderates the relationship between environmental congruence and ILP. For individuals with a low level of CA_TASK, the relationship between environmental congruence and ILP will be higher than for individuals with a high level of CA_TASK.

In the field of Information Systems, CA has been conceptualized as the state of being completely immersed in a technology (CA_TECH, Agarwal and Karahanna 2000; Burton-Jones and Straub 2006), letting the role of the context strongly depend on the technology referred to. For example, in the study of Agarwal and Karahanna (2000), the technology in which an individual was cognitively absorbed in was *the web*, whereas Burton-Jones and Straub (2006) referred to *MS Excel*. If these constructs initially developed in the context of technology acceptance are adapted in the field of learning, it is important to note the different implications a high CA_TECH might have in both cases. If an individual has to learn something in the web, a high CA in the web does not necessarily imply a high CA_TASK because the web can be used in a range of task-unrelated ways very easily. In contrast, MS Excel still can be used in task-unrelated ways while being highly cognitive absorbed (e.g. drawing pictures instead of calculation), but the affordance for these alternatives is probably much lower than in the case of the web.

In the context of IVR, CA_TECH leaves the user even more room for task-unrelated activities. By sealing the participants from the actual world through a head-mounted display and earphones, an almost completely immersing new virtual world is created. Therefore, context that would traditionally be neither part of the task nor part of the technology (e.g., a cupboard displayed in the IVR) becomes a part of the technology. Thus, the meaning of CA_TECH changes dramatically in IVR by covering a much broader range of the environment.

Table 1. Construct definitions

Construct	Definition	Source
Cognitive absorption in task (CA_TASK)	Cognitive absorption is defined as an enjoyable state of deep (cognitive) involvement in the *performed task.*	(Csikszentmihalyi 1990)
Cognitive absorption in technology (CA_TECH)	Cognitivie absorption is defined as an enjoyable state of deep (cognitive) involvement with the *technology used.*	Agarwal and Karahanna (2000)
Telepresence	Telepresence refers to perception of the user in contrast to the technology design. It is defined as the degree to which an individual perceives to be in a distant place.	(Schultze 2010, 2014)
Sensory Immersion	Sensory immersion describes the design of the technology in contrast to the perception of the user. It is defined as the degree to which a technology can achieve convincing illusion of reality to the users' senses.	(Schultze 2010, 2014; Slater and Wilbur 1997)

Whereas the described unspecificity of CA_TECH is not that important for studies of technology acceptance, it needs to be addressed in the area of learning because of the confusion with CA_TASK. Studies that have used items that resembled CA_TECH instead of CA_TASK in the learning context showed that CA_TECH might enhance learning through a motivational route by affecting learner satisfaction (Leong 2011) and continued use (Guo et al. 2016) as well as perceived learning (Reychav and Wu 2015). However, these studies did not vary cognitive absorption experimentally and used technologies such as computers, smartphones or tablets instead of IVR. For the relationship between CA_TECH and learning in an IVR, a qualitative research gives initial support for a relationship between CA_TECH and learning (Kampling 2018). Therefore, we want to address this research gap and investigate whether CA_TECH has an influence on actual (instead of perceived) learning outcomes for declarative knowledge.

In the field of context-dependent learning in an IVR, CA_TECH might influence the relationship between environmental congruence and learning. We assume that a higher CA_TECH before the learning task will lead to a stronger encoding of contextual information which can then strengthen the relationship between environmental congruence and learning.

H3: CA_TECH moderates the relationship between environmental congruence and ILP. For individuals with a high level of CA_TECH, the relationship between environmental congruence and ILP will be stronger than for individuals with a low level of CA_TECH.

Immersive Virtual Realities and Cognitive Absorption

IVR can enhance the sense of "being there" – usually called telepresence – (Schultze 2010, 2014) by presenting a high degree of sensory immersion to the user. Whereas telepresence refers to the psychological perception of the user, sensory immersion refers to the objective criteria of the technology design. Sensory immersion is therefore defined as the degree to which a technology can achieve an inclusive, extensive, surrounding and vivid illusion of reality to the users' senses, matches of the user and matches the users' movements to the visualizations of the IVR, and presents a convincing plot to the senses of the user (Slater and Wilbur 1997).

Different factors of sensory immersion influence telepresence positively (Cummings and Bailenson 2016) which in turn is positively related to CA_TECH (Faiola et al. 2013). High sensory immersion should therefore lead to higher telepresence and CA_TECH than low immersion. We therefore hypothesize an interaction effect of sensory immersion and environmental congruence on ILP, which is mediated by CA_TECH for sensory immersion.

H4a: There is an interaction effect of sensory immersion and environmental congruence on ILP. High sensory immersion strengthens the effect of environmental congruence on ILP more than low sensory immersion.

H4b: The interaction effect of environmental congruence and sensory immersion is mediated by CA_TECH for sensory immersion.

Method

Participants and Design

We plan to recruit 200 students of Information Systems and Business for the experiment who receive a compensation of 5€ for their participation. We use a 2 (learning environment: Room A vs. Room B) x 2 (immersion: low vs. high) x 2 (recall environment: Room A vs. Room B) between subjects design. We use Information Systems and Business students as participants because with them, we can let them take the IVR experience at a physical location where they learn frequently (the IS department). For the recall setting, we can then use one place in which they have never been (the research center) and one place in which they are only occasionally (the main university). By doing this, we create a similarity to the situation in which the user wants to learn in an environment which they can visit only with difficulty and therefore does not necessarily visit it often.

Materials

Hardware and Software. The entire virtual environment is designed with the game engine Unity and the use of a 360° camera as well as the use of 3D laser scanning for the two rooms in which the learning takes place. The use of the 3D laser scanning makes sure that the participants can walk freely in the Room And sit on a chair and at a table that are modeled in accordance with the real ones in both contexts. The chair is tracked with a HTC Vive Tracker to allow participants to sit down without falling. All participants will wear a head-mounted display (HTC Vive) for viewing the IVR in the learning phase. In the front of the HTC Vive, the Leap Motion technology (similar to Schwind et al. 2017) is mounted for all participants, even though it will display the tracked hands, recognized by optical sensors, into the virtual scene in real-time only for participants in the high sensory immersion condition. These participants can then act and interact (touching, moving, manipulating etc.) with objects through a virtual model of their real hands in contrast to using the HTC-Vive controller of the low immersion condition. Additionally, we will let the participants in all conditions wear three HTC Vive trackers (two on each foot and one on the hip) for full-body tracking with Ikinema Orion which are also only functional for participants in the high immersion condition[1]. For the audio aspects within the experiments, a noise cancelling headphone will be used. At the beginning, each participant will be fitted with the headphone and active noise canceling.

Learning Task. Comparable to similar studies used for context-dependent learning (e.g., Godden and Baddeley 1975; Smith et al. 1978), we will use a word list consisting of 40 common, four-letter words

[1] We use this approach to control for effects of the hardware. https://ikinema.com/orion

that the participants have to remember. The words will be presented via headphones and the space between words will be an interval of 3 seconds.

Sensory immersion. In the low sensory immersion condition, participants wear a head-mounted display and use controllers to interact with the virtual environment and wear headphone through which no sound is played. Instead of having a body, participants only see two controllers with which they interact in the virtual world. In the high sensory immersion condition, participants wear a head-mounted display and, using Leap Motion, they can interact with the virtual environment using their hands which are displayed through Leap Motion in the VR. Additionally, they can see a body when they look down which is tracked through the HTC Vive trackers. Background music is played through the headphones, different for each context.

Contexts. Both contexts are presented virtually in the laboratory of the Information Systems department of the local university for the learning phase and are later visited physically for the recall phase. The contexts differ in how the two rooms look and where they are located. Room A is located at the research center of the local university which is about 15 minutes by bus from the Information Systems department. Room A is designed similar to an office and participants sit at a table on which a desktop computer, a telephone and various office tools stand. They look at a flip chart and a cupboard filled with books. A specific background music is played in the room. Room B is located at a building in the main university which takes about 20 minutes by bus from the Information Systems department and about 10 minutes by bus from the research center. The room is associated with a specific background music consisting of different classical music pieces. Room B is similar to an office and participants sit at a table on which a desktop computer, a telephone and various office tools stand. They look at a flip chart and a cupboard filled with books. A different background music consisting of other classical music pieces is played for Room B with the same tonality and tempo (similar to Isarida et al. 2017).

Procedure

The experiment is divided in two sessions, the first one for the learning phase and the second one for recall. The first phase, were the exposition to the IVR-setting takes place, is located at the Information Systems department of the local university. Participants are tested individually. When they enter the laboratory, the experimenter tells them that the experiment is about experiences in VR and explains them how to put on the head mounted-display and the trackers. After participants have put on the head-mounted display, they see the outside of the building in which Room A or Room B is located in a 360° video. Then they see a virtual walk through the door of the building and enter it. In the building, they walk to the door of either Room A or Room B (depending on the building). Participants are then instructed by headphones to open the door either with the controllers in the low immersion condition or with their hands in the high immersion condition. They can then walk freely towards the chair to sit at the table. After they answer the presence and CA_TECH questionnaires, they are told that they will hear a word list and that they should try to remember the words. Participants then hear the 40-word word list. Afterwards, similar to the procedure of Smith et al. (1978) the word list is presented again and participants have ten seconds between each word to rate the affective value of each word on a continuum from “good” to “bad” using either the controllers (low body tracking) or their hands (high body tracking). We use this approach to induce a sense of closure for the session and prevent participants from rehearsing the list between sessions. Participants then answer the questionnaire for CA_TASK. Participants are told that they should come to either Room A or Room B on the next day at a specific time in order to answer a final questionnaire and to receive their compensation fee.

The second session takes place about 24 hours later and is located either at the research center of the university (Room A) or a building in the main university (Room B). When they arrive at the room, the experimenter explains to the participant that they should write down as many words as they can remember in a surprise free recall test. The experimenter then leaves the room for 10 minutes. Afterwards, subjects are asked whether they have rehearsed any words between sessions, fill out the questionnaire of perceived room similarity between learning and recall room, answer questions for perceived learning, receive their compensation fee and are thanked and debriefed.

Measures

Individual learning performance. Individual learning performance is measured by the number of items recalled and by a perceived learning questionnaire adopted from Magni et al (2013).

Cognitive absorption. We adapt the 5-item-measure of Burton-Jones and Straub (2006) for CA_TECH and CA_TASK. We frame the CA_TECH items towards the technology, similar to Burton-Jones and Straub, and the CA_TASK items towards the task, similar to Magni et al. (2013).

Manipulation checks. We use the Igroup Presence Questionnaire (Schubert et al. 2001) as manipulation check for Immersion and questions for perceived room similarity as manipulation checks for environmental congruence.

Data analysis

We will analyze the data using ANOVA and covariance-based structural equation modeling.

Discussion

With the experiment, we plan to show that environmental congruence can be designed in IVR and that is enhanced by sensory immersion through CA_TECH. By letting participants learn a word list, an approach that is often used in basic research on memory, we want to show that the proposed explanatory design theory can be used for a range of different tasks. Whereas the relevance for practice would have been more obvious with a task that focused on application in an organization, recall of a word list represents a basic function of memory in general. Therefore, the underlying mechanisms of the explanatory design theory should apply for most tasks in which recall of declarative knowledge is relevant. On the basis of our results, we expect that future research can extend our explanatory design theory to different types of knowledge (e.g. implicit knowledge), compare it with the effects in reality, and use it as basis for identifying additional design options to enhance ILP.

References

Agarwal, R., and Karahanna, E. 2000. “Time Flies When You’re Having Fun: Cognitive Absorption and Beliefs About Information Technology Usage,” *MIS Quarterly* (24:4), pp. 665–694.

Baskerville, R., and Pries-Heje, J. 2010. “Explanatory Design Theory,” *Business & Information Systems Engineering* (2:5), pp. 271–282.

Burton-Jones, A., and Straub, D. W. 2006. “Reconceptualizing System Usage: An Approach and Empirical Test,” *Information Systems Research* (17:3), pp. 228–246.

Canas, J. J., and Nelson, D. L. 1986. “Recognition and Environmental Context: The Effect of Testing by Phone,” *Bulletin of the Psychonomic Society* (24:6), pp. 407–409.

Csikszentmihalyi, M. 1990. *Flow: The Psychology of Optimal Experience*, New York: Harper & Row.

Cummings, J. J., and Bailenson, J. N. 2016. “How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence,” *Media Psychology* (19:2), pp. 272–309.

Faiola, A., Newlon, C., Pfaff, M., and Smyslova, O. 2013. “Correlating the Effects of Flow and Telepresence in Virtual Worlds: Enhancing Our Understanding of User Behavior in Game-Based Learning,” *Computers in Human Behavior* (29:3), pp. 1113–1121.

Fernández, Á., and Alonso, M. A. 2001. “The relative value of environmental context reinstatement in free recall,” *Psicológica* (22:2).

Godden, D. R., and Baddeley, A. D. 1975. “Context-Dependent Memory in Two Natural Environments: On Land and Underwater,” *British Journal of Psychology* (66:3), pp. 325–331.

Gregor, S. 2009. “Building Theory in the Sciences of the Artificial,” in *Proceedings of the 4th*

International Conference on Design Science Research in Information Systems and Technology, ACM, pp. 1–10.

Guo, Z., Xiao, L., Van Toorn, C., Lai, Y., and Seo, C. 2016. "Promoting Online Learners' Continuance Intention: An Integrated Flow Framework," *Information & Management* (53:2), pp. 279–295.

Hayden, S. 2018. "Volkswagen Group to Train 10,000 Employees in VR This Year," *Road to VR*, , February 28.

Isarida, T., and Isarida, T. K. 2014. "Environmental Context-Dependent Memory," *Advances in Experimental Psychology Research*, pp. 115–151.

Isarida, T., and Isarin, T. K. 2007. "Environmental Context Effects of Background Color in Free Recall," *Memory & Cognition* (35:7), pp. 1620–1629.

Isarida, T. K., Kubota, T., Nakajima, S., and Isarida, T. 2017. "Reexamination of Mood-Mediation Hypothesis of Background-Music-Dependent Effects in Free Recall," *Quarterly Journal of Experimental Psychology (2006)* (70:3), pp. 533–543.

Isarida, T., Sakai, T., Kubota, T., Koga, M., Katayama, Y., and Isarida, T. K. 2014. "Odor-Context Effects in Free Recall after a Short Retention Interval: A New Methodology for Controlling Adaptation," *Memory & Cognition* (42:3), pp. 421–433.

Kampling, H. 2018. "The Role of Immersive Virtual Reality in Individual Learning," in *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Kuechler, B., and Vaishnavi, V. 2012. "A Framework for Theory Development in Design Science Research: Multiple Perspectives," *Journal of the Association for Information Systems* (13:6), pp. 395–423.

Leong, P. 2011. "Role of Social Presence and Cognitive Absorption in Online Learning Environments," *Distance Education* (32:1), pp. 5–28.

Magni, M., Paolino, C., Cappetta, R., and Proserpio, L. 2013. "Diving Too Deep: How Cognitive Absorption and Group Learning Behavior Affect Individual Learning," *Academy of Management Learning & Education* (12:1), pp. 51–69.

Niehaves, B., and Ortbach, K. 2016. "The Inner and the Outer Model in Explanatory Design Theory: The Case of Designing Electronic Feedback Systems," *European Journal of Information Systems* (25:4), pp. 303–316.

Preacher, K. J., Rucker, D. D., and Hayes, A. F. 2007. "Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions," *Multivariate Behavioral Research* (42:1), pp. 185–227.

Reychav, I., and Wu, D. 2015. "Are Your Users Actively Involved? A Cognitive Absorption Perspective in Mobile Training," *Computers in Human Behavior* (44), pp. 335–346.

Schultze, U. 2010. "Embodiment and Presence in Virtual Worlds: A Review," *Journal of Information Technology* (25:4), pp. 434–449.

Schultze, U. 2014. "Performing Embodied Identity in Virtual Worlds," *European Journal of Information Systems* (23:1), pp. 84–95.

Slater, M., and Wilbur, S. 1997. "A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments | Presence: Teleoperators and Virtual Environments | MIT Press Journals," *Presence: Teleoperators and Virtual Environments* (6:6), pp. 603–616.

Smith, S. M., Glenberg, A., and Bjork, R. A. 1978. "Environmental Context and Human Memory," *Memory & Cognition* (6:4), pp. 342–353.

Smith, S. M., and Vela, E. 2001. "Environmental Context-Dependent Memory: A Review and Meta-Analysis," *Psychonomic Bulletin & Review* (8:2), pp. 203–220.