Remembering the Physical As Virtual: Source Confusion and Physical Interaction in Augmented Reality

Ajoy S. Fernandes*

Graduate Student Computer Science
Columbia University

Ranxiao Frances Wang[†]

Professor of Psychology
University of Illinois at
Urbana-Champaign

Daniel J. Simons[‡]

Professor of Psychology
University of Illinois at
Urbana-Champaign

Abstract

This study explored whether people misremember having seen a physical object when they actually had viewed a virtual one in augmented reality (and vice versa). Participants viewed uniquely shaped objects in a virtual form or a physical, 3D-printed form. A camera mounted behind a computer monitor showed either the physical object or an augmented reality version of it on the display. After viewing the full set of objects, participants viewed photographs of each object (taken from the physical version) and judged whether they had originally seen it as a physical or virtual object. On average, participants correctly identified the object format for 60% of the photographs. When participants were allowed to manipulate the physical or virtual object (using a Leap Motion Controller), accuracy increased to 73%. In both cases, participants were biased to remember the objects as having been virtual.

CR Categories: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems —Artificial, augmented, and virtual realities I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality; J.4 [Social and Behavioral Sciences]: Psychology;

Keywords: augmented reality, object recognition, perception and action, human perception and performance

1 Introduction

Both perception and imagination can produce memories, and people regularly confuse the source of those memories, a failure of “reality monitoring” [Johnson and Raye 1981]. Reality monitoring is the process involved in discriminating between memories of real and imagined events. For example, we can remember having witnessed an event when we only read about it [Durso and Johnson 1980]. Technology may amplify such recall failure, with people misremembering virtual events as real. Although people apparently can distinguish memories of virtual events from physical events [Hoffman et al. 1995], early “virtual reality monitoring” studies were limited by the quality of the virtual experience (e.g., field of view, update rate, resolution), making virtual events more distinct from real ones.

Unlike a fully-immersive virtual environment, augmented reality integrates virtual objects into a depicted real environment [Azuma

1997]. This integration of real and virtual may amplify confusion about the sources of our experiences. Will people falsely remember augmented reality objects as if they had been real? Will they confuse real objects for virtually experienced ones?

For virtual reality, situational cues may disambiguate the real from the virtual: the act of wearing a headset, the quality of the graphics, and the emotions and feelings accompanying immersion in a virtual world all provide cues about the source of an experience [Hoffman et al. 1995]. Augmented reality eliminates many of those cues, and improving graphics minimize the difference in the appearance of virtual and real objects, increasing the chances of source confusion. In the absence of any visual differences between physical and virtual objects, people would have to rely on other information to disambiguate virtual and real objects. For example people could infer that a mailbox floating in the air must be virtual even if it were visually indistinguishable from a physical mailbox. Without such contextual information, source confusion would be inevitable.

In the absence of visual differences between physical and virtual objects in augmented reality, other sources of information might mitigate source confusion. One potential source of disambiguating information is action. Interacting with real and virtual objects might enhance memory for their format even in the absence of visual differences. To the extent that our motor actions influence our memories [Garbarini and Adenzato 2004], such actions might provide cues to the source of our experiences.

In this study, we explored whether or not people misremember augmented reality objects as real ones (and vice versa), and we examined whether manipulating the virtual and real objects at study yields better memory for their format.

2 Method

2.1 Objects

To measure whether people confuse virtual and real objects, we first needed to develop a set of unfamiliar objects that participants could distinguish and remember. To do so, we used an iterative process in which we created virtual objects and tested whether participants could differentiate them. For each such test, participants viewed a set of 30 virtual objects and then were asked to view a randomly ordered sequence of 60 subject objects (half new, half studied). For each, they indicated whether or not it was one they had studied. Whenever participants confused an old object for a new one (or vice versa), we eliminated one of the two objects and added a new object. After this informal testing with a total of 19 participants, we developed a set of 60 distinctive objects. Our primary experiment used 58 of these objects (2 failed to print correctly).

For each of these objects, we created both a 3D printed form and a virtual graphic, and we took a photograph of the printed object. The physical and virtual forms were designed to have equal size, color (red), and shading when viewed on the computer monitor. Note, though, that the virtual and physical objects were visually distinguishable based on differences in shading and shadowing.

*e-mail: asf2168@columbia.edu

[†]e-mail: wang18@illinois.edu

[‡]e-mail: dsimons@illinois.edu

Figure 1: Screenshot of a Virtual (Left) versus Physical (Right) Object as Displayed on the Monitor.

Figure 2: The 58 Objects Used.

2.2 Participants and Procedures

Eighty-four undergraduate students participated in exchange for extra credit in their psychology courses. The participant sat at a large table, separated from the experimenter by a curtain divider. The arrangement allowed the experimenter to place objects on the participants table without their actions being seen. The participants view of their table was almost completely occluded by a 15" Dell LCD monitor rotated 90 degrees to portrait mode.

The monitor displayed the real-time output of a web camera focused on the table surface directly behind the monitor. The participant responded using a keyboard positioned on their lap. A Leap Motion 3D Controller was placed to the side of the monitor. Software used to control the experiment was designed and implemented in Unity 3D.

Participants were randomly assigned to one of two experimental conditions. In the "Saw" condition (N=37), participants only viewed the objects whereas in the "Manipulated" condition (N=47), they handled the objects either directly or virtually (using the Leap Motion 3D Controller).

In both conditions, participants first encountered the objects in a "study" phase and then were tested on their memory for the format of the studied objects. Prior to the study phase, participants viewed a virtual and physical object on their monitor and they were told that they would view an equal number of virtual and physical objects.

During the study phase, half of the 58 objects were presented phys-

Figure 3: Top View Showing Layout of Experiment Room.

ically and half were presented virtually. Which objects were presented in each format was randomly determined for each participant, as was the order of presentation for the 58 objects. The computer display was blank while the experimenter prepared each trial. For "physical" trials, the experimenter placed the object on a marked location on the table so that it would be visible to the web camera. For the "virtual" trials, the computer automatically placed the object in the appropriate location for the augmented reality display so that it matched the position, lighting, and orientation of the corresponding physical version of that object. On virtual trials, the experimenter placed a physical object on the mark on the table and then removed it silently to ensure that participants could not distinguish physical and virtual trials based on the experimenters actions or auditory cues. Once the object was placed, the participant was instructed to press "enter" on their keyboard and the display appeared on their monitor.

In the "Saw" condition, the participant simply viewed the object on the monitor for as long as they wished. They were asked to press the left arrow key if they thought the object was physically present and the right arrow key if they thought it was virtual. In the "Manipulated" condition, participants physically picked up and handled the object behind the monitor. Or, when the object was virtual, they manipulated it via the Leap controller. For the virtual condition, the object was suspended slightly above the table surface, and the Leap interaction was mapped such that the rotation of the virtual object corresponded to the rotation of an open palm facing the floor. The Leap was placed to the side of the object so that the object displayed in the monitor, would not be occluded by the participants hand. As in the "Saw" condition, after interacting with the object, participants used the arrow keys to indicate whether it was physical or virtual. After the keyboard response, the screen went blank while the experimenter prepared the next trial.

After the study phase, a text screen appeared indicating the end of the study phase and the start of the test phase. During the test phase, participants were tested on their memory for the format of each studied object. On each trial, a photograph of each object appeared as it would have in the physical condition and participants used the left and right arrow keys to indicate whether they had studied the object in a physical or virtual form. The 58 objects were presented in a different random order for each participant.

3 Results and Analysis

During the study phase of the experiment, participants reliably differentiated physical from virtual objects, even when they only saw

Table 1: Percentage accuracy and standard deviations by condition for each outcome measure, along with Bayes Factors, t-tests, and effect sizes for the comparison between the conditions. Bayes Factors were computed using JASP, with a Cauchy prior width of 0.707. Effect size and t-tests are based on a between-subjects comparison of the "Manipulated" and "Saw" conditions for that measure.

Measure	Mean % (SD)	Tests
Accuracy		
Manipulated	73.3 (10.6)	BF = 474036.9, t(82) = 6.20, p < .001, d = 1.36
Saw	60.4 (7.7)	
Virtual Objects		
Manipulated	75.4 (11.9)	BF = 14.0, t(82) = 3.12, p = .002, d = 0.69
Saw	66.4 (14.5)	
Physical Objects		
Manipulated	71.3 (12.2)	BF = 101124.0, t(82) = 5.81, p < .001, d = 1.28
Saw	54.5 (14.2)	
Bias to Respond “Virtual”		
Manipulated	52.0 (5.7)	BF = 1.18, t(82) = 1.95, p = .055, d = 0.43
Saw	55.9 (12.1)	

the objects ("Saw": $M=95\%$ correct; "Manipulated": $M=98\%$). Accurate performance during the study phase of the "Saw" condition means that the objects were visually discriminable and that physical manipulation was not necessary to encode objects as physical or virtual.

In both conditions participants correctly remembered the form of objects better than would be expected by chance (one sample $t(36)=8.26$, $p < .001$, $d=1.35$ for the "Saw" condition; $t(46)=15.03$, $p < .001$, $d=2.19$ for the "Manipulated" condition). At test, participants who manipulated the objects at study were better able to remember their format than were participants who only saw the objects (See Table for means and statistical comparisons). That pattern, better performance for the "Manipulated" group than the "Saw" group, held true for both virtual and physical studied objects. Participants in both groups were more likely to respond virtual at test (one sample $t(36)=2.97$, $p < .01$, $d=.49$ for the "Saw" condition; $t(46)=2.43$, $p < .05$, $d=.35$ for the "Manipulated" condition), but this bias did not differ substantially across conditions (the Bayes Factor provides only anecdotal evidence for a difference in bias between the conditions).

4 Discussion

Increasing the perceptual similarity between memories of real and imagined events tends to increase reality monitoring confusions [Johnson et al. 1979]. Augmented reality eliminates many of the external cues that differentiate more traditional virtual reality from physical reality (e.g., the headset). Consequently, we should expect source monitoring confusion even for virtual objects that are visually discriminable from physical ones.

Participants in our task could reliably discriminate virtual and physical objects viewed via a web camera on a computer monitor, even

when they were not allowed to manipulate the objects. Yet, when asked to remember whether they had studied objects in virtual or physical form, they did so with greater accuracy when they had manipulated the objects at study. When they simply viewed the objects at study, they often misremembered physical objects as having been virtual, and vice versa.

There are several potential explanations for better performance when participants physically manipulated the objects. First, when physically manipulating the objects, participants observed multiple views of the objects. Second, manipulating the objects might have increased the relative engagement in the task, leading participants to pay more attention to the objects. Third, exerting active control over changes in view can enhance memory [Simons et al. 2002].

Given that participants could achieve 50% accuracy with no memory for the object source (by guessing), to determine how much people remembered, we need to take guessing into account. To obtain an overall accuracy of 73.3%, such as in the Manipulated condition, participants would need to remember the format of 46.6% of the studied objects while guessing randomly on the remaining trials (half of their guesses would be correct). To achieve 60.4% accuracy in the Saw condition, participants would need to remember the object format on 20.8% of trials and guess accurately on 50% of the remaining trials. Thus, even though physically manipulating the objects led to less source confusion, participants in our study still appeared to confuse the source format of the studied object more than 50% of the time. Following similar reasoning, when they only saw the objects, they remembered the correct form of the objects on about 20% of the trials.

5 General Discussion

Participants consistently confused the source of studied objects, mistaking physically studied objects for virtual ones. This study provides insights into how future augmented reality technologies could affect our memories of our daily experiences. Even though the virtual objects were visually distinguishable from physical objects, participants still confused the source of their memories. As the realism of augmented reality graphics increases, such source confusion is likely to increase as well.

Several design choices limit the conclusions we can draw from this study. First, we did not control observation time across conditions—we chose instead to allow participants to study the objects as long as they wished. We chose this approach to avoid inducing source memory failures due to time constraints on encoding. Consequently, the encoding time was not strictly equated across conditions, and that difference in study time might introduce different source memory in the two conditions. That difference could be amplified if participants were more engaged in the manipulation condition. Future studies could systematically control encoding time to observe the effect on differences in source memory across conditions.

Second, given time constraints on the testing sessions, each participant could complete a limited number of trials. We chose to compare the Manipulated and "Saw" conditions across participants in order to increase the precision of the accuracy dependent measure for each participant (by maximizing the number of trials contributing to that estimate). Using a between-groups design relies on random assignment to eliminate contributions of individual differences to the group differences. Future studies could employ a within-participant design to reduce the contribution of such differences, thereby increasing the power to observe differences between groups.

Third, the source confusion we documented occurred when partici-

pants were tested on their memory almost immediately after studying the objects. Any additional delay between learning and testing likely would amplify failures of source memory. That said, although the objects in our study were distinguishable, they also were arbitrary and similar in many respects (e.g., red, meaningless, etc). Perhaps source confusion would be lessened for meaningful objects embedded in more naturalistic augmented reality contexts.

Conclusion

In our study, the virtual and physical objects were differentiable by sight alone, but eventually they will not be, and source confusion will be unavoidable. Source confusion could be reduced by providing other information about the format of the displayed objects, if desired, but even when participants physically manipulate the virtual or real objects, they likely will misremember whether the objects appeared virtually or in reality.

References

- AZUMA, R. 1997. A survey of augmented reality. *Presence* 6, 4, 355–385.
- DURSO, F., AND JOHNSON, M. 1980. The effects of orienting tasks on recognition, recall, and modality confusion of pictures and words. *Journal of Verbal Learning and Verbal Behavior* 19, 4, 416–429.
- GARBARINI, F., AND ADENZATO, M. 2004. At the root of embodied cognition: Cognitive science meets neurophysiology. *Brain and Cognition* 56, 1, 100–106.
- HOFFMAN, H., HULLFISH, K., AND HOUSTON, S. 1995. Virtual-reality monitoring. In *Proceedings Virtual Reality Annual International Symposium*.
- JOHNSON, M., AND RAYE, C. 1981. Reality monitoring. *Psychological Review* 88, 1, 67–85.
- JOHNSON, M., RAYE, C., WANG, A., AND TAYLOR, T. 1979. Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory* 5, 3, 229–240.
- SIMONS, D., WANG, R., AND RODDENBERRY, D. 2002. Object recognition is mediated by extraretinal information. *Perception and Psychophysics* 64, 4, 521–530.