Visual Behavior During Engagement with Tangible and Virtual Representations of Archaeological Artifacts

Niveta Ramkumar
University of New Hampshire
Durham, New Hampshire
nr1063@wildcats.unh.edu

Nadia Fereydooni
University of New Hampshire
Durham, New Hampshire
nf1021@wildcats.unh.edu

Orit Shaer
Wellesley College
Wellesley, Massachusetts
oshaer@wellesley.edu

Andrew L. Kun
University of New Hampshire
Durham, New Hampshire
andrew.kun@unh.edu

ABSTRACT

In this paper, we present results from a study of users' visual behavior while engaging with tangible and virtual representations of archaeological artifacts. We replicated and extended a recent study that introduced an augmented reality system implemented using HoloLens, for engaging with the artifacts. Our study goes beyond the original study to estimate the distribution of users' visual attention for both tangible and virtual representations of the artifacts. Our study confirmed the results of the original study in various aspects. Specifically, participants in both studies confirmed the immersive nature of the HoloLens condition and showed similar learning outcomes in terms of post-task open questions. Additionally, our findings indicate that users allocate their visual attention in similar ways when interacting with virtual and tangible learning material, in terms of total gaze duration, gaze on object duration, and object fixation duration.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; *Gestural input*; • **Computing methodologies** → Object recognition.

KEYWORDS

Human-Centered Computing, Mixed/Augmented reality, Gesture input, Object recognition, Eye tracking, Object-based learning

ACM Reference Format:
Niveta Ramkumar, Nadia Fereydooni, Orit Shaer, and Andrew L. Kun. 2019. Visual Behavior During Engagement with Tangible and Virtual Representations of Archaeological Artifacts. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays (PerDis '19), June 12–14, 2019, Palermo, Italy.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3321335.3324930

PerDis '19, June 12–14, 2019, Palermo, Italy

ACM ISBN 978-1-4503-6751-6/19/06…$15.00
https://doi.org/10.1145/3321335.3324930

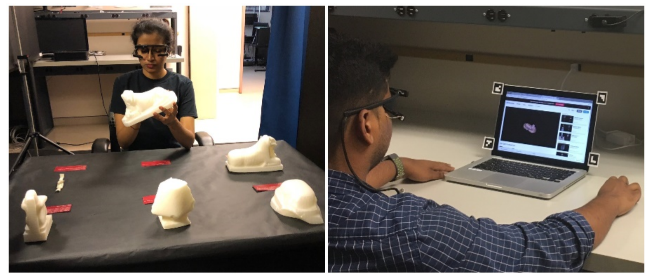

Figure 1: Inventory of 3D artifacts and the participant exploring their chosen artifact (left). A participant wearing eye tracker working with the Sketchfab platform (right)

1 INTRODUCTION

Object-based learning emphasizes the student's interaction with physical artifacts in the learning process. This pedagogical approach has been found to be more effective than relying exclusively on lectures[7]. This approach is well established in various fields including archaeology, art history, and anthropology[33]. The advent of technologies such as virtual reality (VR), augmented reality (AR), and 3D fabrication, has created opportunities for implementing object-based learning without the need to access the original physical artifacts[8]. These technologies allow educators to create tactile and virtual models of the artifacts so that students can learn by exploring and analyzing these models [30]. However, relatively little is known about how tactile and virtual models can be used in object-based learning. Recently, Pollalis et al.[30] conducted an experiment to evaluate learning with three different representations of ancient artifacts. Users interacted with artifacts represented as 3D models on a computer screen, as 3D virtual models in augmented reality, and as 3D fabricated tangible objects. Pollalis et al. found that there were differences in learning outcomes for the three types of presentations. The study we present here replicates core aspects of this study but extend it by asking the following two questions. First, how does visual attention vary among the three conditions? Second, how are differences in visual attention related to the differences in learning outcomes?

Replicating studies in HCI is important because practitioners and researchers could better trust and build upon results from the studies of novel technologies that can be, and have been, replicated.

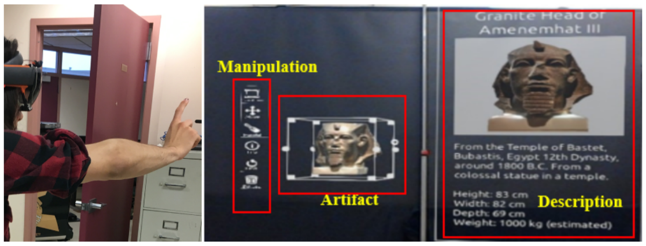

Figure 2: Participant wearing HoloLens with the add-on eye tracker(right). Screenshot of the hologram indicating the three areas of interest (AOI) in the Hololens condition(left).

"Replicate and extend" studies in particular, test the limits and pertinence of previous results[6, 11]. In this study, we replicate the tasks and experimental designs utilized by Pollalis et al. while extending it by collecting eye tracking data to study the visual behavior of participants. By observing visual behaviors, we can provide a quantitative measure of users' interaction with the artifacts, which can in turn help us understand the reasons for the observed learning outcomes. This understanding can then assist us in improving existing and future interactive learning tools. Here we report the findings form a study comparing how users interact with three representations of objects: 3D printed replicas of museum artifacts, holographic replicas, and 3D digital models displayed on a screen. We used an eye tracker to follow users' gaze movements to study their visual behavior. Our contributions beyond replicating the results of the original study include: 1) a new study using eye tracking to analyze users' visual behavior while learning about 3D objects; 2) computational methods for analyzing visual behavior around 3D objects and digital 3D digital models; and 3) understanding of users' visual behavior and how it is related to learning outcomes in an object-based learning activity.

2 BACKGROUND AND RELATED WORK

2.1 Object-Based Learning

Object based learning pedagogy views the learner's interaction with objects as critical for the learning process. Direct interaction with objects allows learners to take charge of their learning process and construct meanings to enhance their critical thinking skills[15]. Much research indicates the benefits of AR for learning and problem solving [3, 18, 27]. AR has been shown to be useful in motivating students in the learning process. Taking advantage of the features of this technology allows educators to improve students' educational experience, their engagement, and their academic achievement [2, 3]. On the other end, evidence shows that concrete visual models, such as 3D printed replicas, could capture students' attention and provide physical context in which to think about concepts. This makes students feel more comfortable visualizing and describing the material [4, 13].

2.1.1 The original study. Pollalis et al. conducted experiments to understand the learning outcomes when users are engaged with three kinds of replicas of museum artifacts: tangible 3D printed artifacts, 3D virtual models presented on a screen, and holographic artifacts [31]. These specific types of objects on the tangible-virtual spectrum were chosen due to their increasing availability in higher education[14]. The authors assessed users' enjoyment, perceived task workload, spatial presence, and learning outcomes. They observed that object-based learning goals were accomplished comparably with holographic artifacts and with the digital 3D models, while 3D replicas lacked visual information impeding learners' contextualization and critical thinking.

However, further studies are needed to understand how these technologies impact object-based learning processes. Moreover, the roles the physicals and digital elements of the learning experience remain to be mapped out [1]. In this study, we replicated the original experiment by Pollalis et al. [31], while adding eye tracking to each condition. Our goal is to build an understanding of users' visual behavior and how it is related to the learning outcomes reported in the original study.

2.2 Learning and Visual Behavior

Several studies use eye movements to describe users' visual attention, as they are considered the behavioral interface between attention and gaining information from the surrounding environment [17, 32, 40] When considering the learning process and its outcomes, many studies use eye tracking to track how learners interact with the learning material [28], predict their level of comprehension [20], and their learning efficiency [7]. Daraghmi et al. developed an on-screen learning system using eye tracking to give learners feedback about their learning [5]. However, the majority of literature that addresses users' visual behavior in learning focuses on learning material that is presented on a 2D surface [7, 21, 23, 28, 40, 42].

Van der Meulen et al. developed a method to combine eye tracking data with head-tracking data provided by HoloLens in order to improve our ability to assess the gaze location of HoloLens users. We are using this method in our study. However, the AR targets in their study were 2D[39]. In this study, we evaluate visual behavior of learners while learning about artifacts replicated in three different and increasingly available methods of creation: 3D printed physical objects, on-screen digital 3D models, and AR visualization of 3D holograms. To our knowledge, this is the first study using eye tracking to analyze users' visual behavior while learning about 3D objects with these modalities.

3 STUDY

3.1 Experimental Design and Tasks

We conducted a between-subjects experiment in which three groups of participants completed the same learning task. The learning task was developed by Pollalis et al. [31]. It is a task that students might encounter in an archaeology class, and its aim is to enhance students' observational skills and their critical analysis skills. The task consists of selecting two artifacts from an available inventory of six artifacts, exploring them, and answering the corresponding artifact questionnaire. For each object, participants were asked to indicate the first detail they noticed, all the details they observed, and what characteristic about the object made it unique or similar to the other artifacts in the set. We did not impose a time limit on the task. Participants completed the learning task either using tangible 3D replicas, virtual 3D replicas, or holographic objects [29].

Across the 3 conditions we used different replicas of the same 6 archaeological artifacts.

3.1.1 Tangible 3D replicas (3D prints). In the tangible 3D printed condition, subjects could choose an artifact from a 3D printed gallery of six objects. These objects are the same 3D printed artifacts used in the original study [31]. The Printed descriptions next to the artifacts are identical to the ones in the original study (Figure 1). Participants were free to choose and explore the replicas while wearing an eye tracker.

3.1.2 Sketchfab Condition. Participants in the second condition used Sketchfab, an online 3D modeling platform on a desktop[29]. They could choose any of the 3D models in an inventory of six objects, manipulate the chosen artifact using a mouse, and read more about it in a section below it. The platform was configured in a similar fashion to Pollalis et al. [31]. In the original study, the description section was right next to the artifact section, so the participant did not have to scroll down to view the descriptions. However, in the new version of the Sketchfab which we used the user has to scroll down the page to read the description.

3.1.3 Holographic objects (HoloMuse). The third condition consisted of participants using HoloMuse [30], an AR application on Microsoft HoloLens, which was developed and used by Pollalis et. al[31]. It introduces subjects to an inventory of six holographic objects. They could use air gestures to pick and handle the artifacts by moving, scaling and rotating them (see Figure 2). Users were also able to remove the artifact's material to view its surface and reveal supplementary information about the artifact.

3.1.4 Eye Tracking. We used Pupil Labs head-mounted eye trackers [16] to track users' gaze during the experiment. This eye tracker has a world camera capturing the users' environment, and two slide cameras for users' pupils. It then calculates users' gaze based on their pupil movements and maps it onto the video from the world camera to display the target the user is looking at. Participants who interacted with 3D artifact and Sketchfab wore the eye tracker. In the HoloLens condition we used an eye tracker add-on[39]. The inventory of artifacts, their order in the inventory, their descriptions, and the eye tracking method were consistent throughout all three conditions.

3.2 Participants

We collected data from 35 participants (10 female, average age = 23.5, SD=3.2); 12 participants in 3D prints condition (3 female), 10 participants for Sketchfab condition (3 female), and 13 participants for HoloLens condition (4 female). All the participants were given a $10 gift card at the end of the experiment. We dismissed data from 2 participants in the 3D prints condition and 3 participants in the HoloLens condition due to a low eye tracker confidence (<70%). The low confidence resulted from a suboptimal angle of the eye tracker with respect to the subject's pupils. Thus, we report on data from 30 participants (10 female).

3.3 Procedure

After signing the consent forms, the participants filled out a pre-task questionnaire stating prior practice with visual analysis (e.g. art history class), and specifying former experience with 3D modeling software, AR, or VR. Depending on the condition they were randomly assigned to, participants were asked to wear the eye tracker or the HoloLens, were shown an inventory of six artifacts, and were given a brief training on how to choose an object and handle it. Before starting the task, the worn eye tracker was calibrated using screen marker calibration [34]. The HoloLens condition included an additional step in which we connected the HoloLens and its eye tracker to a server computer to synchronize time on both devices, so that we get real time data about the user: their position, head rotation, name of the hologram they are viewing, and their gaze information. In order to minimize the movement of the headset, we ensured that the HoloLens's headband was secured on the user's head. HoloLens condition participants were trained on how to use the device as well. Following the initial stage of the study, participants were given the task of choosing and studying two artifacts. They were asked to fill out an object questionnaire for each chosen object using a laptop we provided. After finishing the task, they were given a post-task questionnaire to fill. This form consisted of 15 questions, each being a 5-point Likert-type ratings ranging from "Strongly Disagree" to "Strongly Agree". A NASA TLX questionnaire [10] and four open-ended questions were also part of the post-task form. HoloLens users were also asked if they experienced any discomfort while performing the task. Collected data includes: questionnaire responses and eye tracking; for the HoloLens condition we logged data from the server and recorded videos from the HoloLens camera using its online portal.

4 DATA ANALYSIS

We used JMP Pro 14 for the statistical analysis of the results. The collected data was initially tested for normality using Shapiro-Wilk test. For the normally distributed data we used ANOVA for mean comparison, t-test and Tukey test for post hoc analysis. For the non-normally distributed data we used non-parametric Kruskal-Wallis test and the Wilcoxon test for post-hoc analysis. To analyze the open-ended questions, we used the same coding-scheme and process as the original study. Our participants were free to interact with the artifacts and manipulate them with no time limit or any restrictions on how to study the artifact. To better understand the distribution of users' visual attention, we defined three areas of interest(AOI): the artifact, the description (where they could read more about the artifact), the manipulation (which they could use to manipulate the artifact). Any visual target other than these three were categorized as "other surfaces" and will focus on the main three AOIs for statistical analysis. We developed algorithms using MATLAB (discussed below) to identify participants' visual targets over the time.

4.1 3D Prints Condition

Initially the image from the world camera (see Figure 3) is converted into grayscale image. The resulting image then goes through binary conversion using thresholding, and noise removal. Every frame goes through this two-stage processing to detect objects and descriptions separately. Subsequently, we developed an algorithm to mark a perimeter around each item using its centroid, as shown by blue stars in Figure 3. In this figure, the green star represents

Figure 3: A frame from the eye tracker's world camera with the artifact (white) and the description (red)(right). (b) The same image after being processed to identify the artifact and description AOI(left)

the gaze location of the participant at the artifact. If the gaze by a participant was inside the box around an item, we concluded that they were looking at that item. Otherwise, we marked that gaze as "other surfaces" as it was not aimed at a place of interest (artifact or description). Unlike the other two conditions, there is no manipulation AOI for the 3D prints; participants use their hands to directly manipulate the tactile objects. We validated the above algorithm by visually inspecting 1000 randomly selected frames, we found that the accuracy of the detection algorithm was greater than 95%.

4.2 Virtual 3D replicas (Sketchfab)

Similar to the 3D condition, the video of the world camera was analyzed frame by frame to identify the three AOI. Markers were used to mark the boundaries of the laptop screen (see figure 4). Our algorithm first filtered out the regions on the screen that were not of interest (using the markers) and separated the scroll bar which is the manipulation AOI. Then the remaining section of the screen undergoes binary conversion and noise removal to identify the artifact/description AOI based on their color; artifacts' background is black, and the description background is white (see figure 4).

4.3 HoloLens Condition

The eye tracker add-on camera could not see the holograms displayed by the HoloLens. Similar to Van der Meulen et al. work, we developed an algorithm to map users' gaze position into the holographic environment [39]. We collected the users' gaze information from the eye tracker, and holographic environment details from HoloLens. We logged information from a server that synchronizes the eye tracker and the HoloLens. Then we combined this information to find the gaze target of the participant in the augmented space; the holographic artifact (with its name), description, manipulations (move, rotate, etc.), or other surfaces.

4.4 Measures and Indicators

4.4.1 Time on Task. Using the timestamp of each data point recorded by the eye tracker, we calculated the time spent exploring each artifact and other AOI for all participants. We used this measure to assess meaningful engagement and determine how it is affected by the interaction styles.

4.4.2 Fixations. We used fixations to evaluate the visual behavior of the participants as they attended to different AOI (e.g. artifacts and descriptions). Fixations are the state of maintaining the gaze at

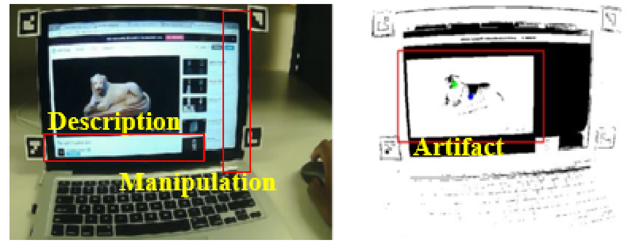

Figure 4: A frame from the eye tracker's world camera (right) and the same image after being processed to identify the artifact and descriptions (left) The red box identifies the artifact region and the green dot shows the user's gaze

a target for a specific amount of time. We extracted fixations with a minimum duration 100 ms[12, 19, 29] and 1°of dispersion angle [25]. We modified the above-mentioned gaze algorithms to find the fixation target. Similar to the gaze AOI, the fixation targets fell into the categories of artifact, manipulation, description, and other surfaces. We explored fixation in terms of fixation rate (fixation count per minute) and duration on the above-mentioned AOI and its overall values.

4.4.3 Learning Outcomes. Study participants were expected to write down detailed explanations of the viewed artifacts. Content codes were used to demonstrate progress from observation to analysis, and to develop a preliminary review of learning outcomes. We used the codes developed by Pollalis et al.[31] since the questionnaire used in our study was the same as the one they composed. The content code are : texture, color, detail, facial feature, damage, material, weight, size, analysis and context. The first two authors acted as coders identifying the content codes in the questionnaire responses. Their inter-code reliability >95%. Disagreements were resolved by consensus.

4.4.4 Perceived Task Workload and Spatial Presence. We used the NASA TLX questionnaire [10][10] used in the original study to measure participants' perceived workload. Another series of questions used in the original study was utilized to measure participants perceived spatial presence. These questions were roughly based on the MEC- SPQ standardized questionnaire [41].

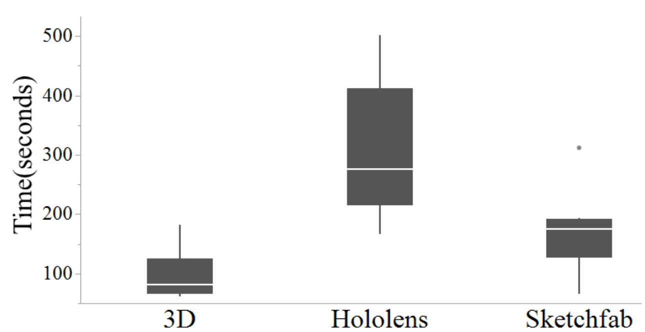

Figure 5: Time participants spent on the task for each condition. The white bar indicates the average time.

5 RESULTS

5.1 Time on Task

The total time participants spent to complete the task can be found in Figure 5. Kruskal-Wallis test indicated that there was a significant effect of condition on total time spent on task [X^2(2)=18.7357, p<0.0001]. The Post hoc comparison showed that the time spent to complete the HoloLens condition task was significantly higher than the other two conditions. This matches the results from the original study for the time on task. Percentage of time distribution among the AOI can be found in Figure 6. There was no significant difference between the time spent on the artifact itself among the three conditions [X^2(2)=3.3626, p=0.1861]. However, the time spent on the description of the artifact differ significantly based on the condition [X^2(2)=13.0477,p=0.0015]. Post hoc testing identified that the time spent on the description was significantly higher in the Sketchfab condition than in the other two conditions. There was a significant difference in the time spent on other surfaces [X^2(2)=19.3652,p<0.0001] and the post hoc analysis indicated significantly higher time spent in the HoloLens condition. Given that there was no manipulation AOI for 3D prints condition, we used the Wilcoxon Test to compare the time that participants spent looking at manipulation in the HoloLens and Sketchfab conditions. We found that the time spent on manipulation was significantly higher in HoloLens condition than Sketchfab condition [Z=14.2857,p<0.0002].

5.2 Fixations

We analyzed fixations from two perspectives; fixation duration, and the fixation rate.

5.2.1 Fixation Rate. The fixation rate (number of fixations per minute) for each condition can be found in Figure 7. The overall fixation rate for the task was not significantly different among the conditions [X^2(2)=1.3871, p=0.4998]. For distribution of fixations among various AOI we report on 28 participants due to momentary server failure for 2 participants in the HoloLens condition (time mismatch). The fixation rate on the artifact was significantly different [X^2(2)=10.8891, p=0.0043]. Post hoc analysis indicated that fixation rate for the HoloLens condition is lower than the other two conditions. Also, the fixation rate on the description was significantly different overall [X^2(2)=19.4813, p<0.0001]. The post hoc analysis indicated that the 3D condition has higher fixation rate on the description than the other two conditions. Participants also had significantly higher fixation rates on manipulation for the HoloLens condition than the Sketchfab condition [Z =8.9195, p=0.0028].

5.2.2 Fixation Duration. We found no evidence that the fixation duration on the artifact [X^2(2)=2.2405, p=0.3262] or description [X^2(2)=14.6945, p=0.0191] were different between the three conditions. However, the t-test results showed that the duration of fixations on the manipulation AOI was significantly higher for the HoloLens condition that the Sketchfab condition [Z=5.4896, p=<0.0006]. Note that there is no separate manipulation AOI for the 3D printed condition.

5.3 Learning Outcomes

We evaluated learning outcomes by counting the number of content codes in the responses to the question asking participants to write down the details they noticed while interacting with the artifact. We found no evidence that the total number of content codes appearing in the responses were different between the conditions [F (2,27) = 1.552, p= 0.8570]. On performing ANOVA on the frequency of mentioning of the individual content codes, we observed that the facial feature and detail code of the visual observation category were significantly different among the conditions. Post hoc testing indicated facial feature was mentioned significantly more often with 3D prints than the other two conditions. Detail was also mentioned significantly less with HoloLens than in the 3D prints and Sketchfab conditions. Rest of the content codes did not show significant difference.

5.4 Perceived Workload and Spatial Presence

There was a significant difference in the perceived workload between conditions [F (2, 27) = 5.7838, p= 0.0081]. Post hoc test suggested that the participants in the 3D prints condition experienced significantly higher effort. Users in the HoloLens condition felt as if the original artifact was physically present in their environment significantly more than users in the other two conditions [X^2(2)=6.3786, p=0.0412]. Participants also claimed to think more intensely about the characteristics of the 3D printed artifacts [X^2(2)=6.7055, p=0.0350].

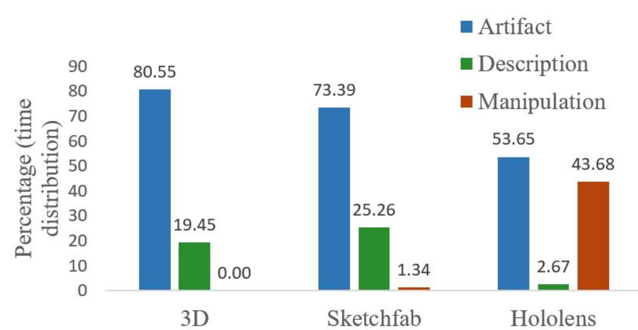

Figure 6: Time percentage distribution for users' gaze among the three areas of interest(AOI) for each condition.

6 DISCUSSION

6.1 Comparison with The Original Study

The time measured in the original study was the overall time spent on the task. Using eye tracking, we were able to classify the total time spent to multiple categories. This enabled us to draw a clear picture of how users distributed their time between analyzing the object, reading about it, and manipulating it. Figure 6 and 7 show this distribution for all three conditions. Our results confirm the results of the original paper in multiple aspects. Both studies found no significant difference in the complexity score of the open question responses about the artifact they viewed. Another common finding was that participants of both studies confirmed the immersive nature of AR by ranking it the highest when asked if they "felt as though the original ancient artifact was physically present in [their] environment". This is despite the fact that 3D printed artifacts were the only ones to include a sense of touch and to exist physically in

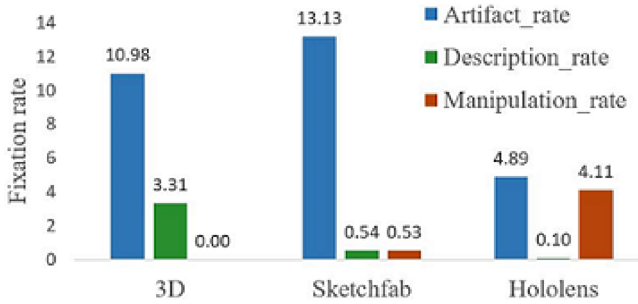

Figure 7: Fixations per minute for each condition.

the environment for the learners. This result suggests a new line of investigation, reconsidering Montessori's finding emphasizing the role of physical material in learning[24]. The total complexity score of the open question responses was comparable among conditions for both studies, However, there were differences with the original study in the particular thematic codes appeared in the responses. The original paper showed that the 3D printed condition was significantly lower in mentions of color, material, and context. Our study did not show such differences among conditions. The reason for this result is not immediately forthcoming; one possibility is difference between participant populations across the two studies, as the studies were conducted at different institutions. Fixations indicate maintaining gaze at a gaze target, during which almost all the visual information is collected [9, 35, 37]. The rate of fixations on the artifact was lowest for the HoloLens condition while the fixation duration was comparable among conditions. Both of the studies had significantly fewer mentions of facial features in the open question responses for the HoloLens condition. It is likely that both the fewer mentions of facial features, and the lower fixation rate are related to two innate features of the holograms: they have lower resolution compared to the Sketchfab condition and are not tactile like the 3D prints condition.

6.2 Physical Artifact vs. Virtual Artifact

Prior research indicates that users have different psychological responses to virtual and tangible objects[1]. However, studies comparing visual behavior for physical and virtual versions of the same task are surprisingly uncommon [22]. One of our contributions is making this comparison. Users in the HoloLens condition spent significantly more time completing the task, however, the total time they spent looking at the object and their object fixation duration was comparable to the other conditions. This is an indication that users' visual behavior towards virtual learning material is similar to tangible ones. This result supports users' claim about the immersive nature of the HoloLens condition; in other words, it seems that participants appreciated the virtual artifacts like the physical ones.

6.3 Interface and Design Implications

Participants reported comparable satisfaction for tangible and virtual interaction types despite technology limitations for current AR equipment. Such limitations like the low resolution, narrow filed of view, and the novelty of the equipment (even though they were trained on how to use HoloLens) posed interaction constraints which participants had to overcome in order to interact with the artifacts. Thus, we are likely to see higher satisfaction measures for the HoloLens condition if interaction becomes more seamless in future products. For example, we observed that the clicking gesture in the AR environment is easier for the users to perform than the dragging gesture. In fact, by visually inspecting the HoloLens videos, we observed that the higher gaze time and fixation duration for manipulation in the HoloLens condition were due to participants' difficulty performing the dragging gesture. Thus, one possible improvement to the interface could be to change how users rotate artifacts: instead of using the dragging gesture, they might prefer to click on a bar that controls object rotation.

Research has shown that interaction costs can lead to increased reflection on the material[9, 26, 36, 38]. Moreover, Marshall claims that the easy manipulation of concrete objects can result in decreased reflection on the learning material [22]. As discussed, the HoloLens condition introduced interaction constraint to the participants. However, in the 3D prints condition users reported that they "thought intensely about the characteristics of the ancient artifact" considerably higher than the other conditions. Based on Marshall's work[22], we expected that participants would report higher thought intensity in the HoloLens condition than in the 3D prints condition, where manipulation was the easiest. This implies that the manipulation effort required in the HoloLens condition might have been too high; this implication is also supported by the fact that participants spent the most time gazing at the manipulation AOI in this condition. Although manipulation for the Sketchfab condition was more complicated than handling the 3D printed artifacts, the Sketchfab condition did not provide the immersive experience for the users which might have been needed

7 LIMITATIONS

One limitation of our detection algorithm is that a minimum color contrast has to be maintained between the 3D printed object and background. However, in the future, we anticipate that 3D printed objects will include color, Hence, future work includes the use of machine learning algorithms to identify various AOI. Although a rare occurrence in our study, another challenge was eye trackers heating up after long usage, creating discomfort for users.

8 CONCLUSION AND FUTURE WORK

By replicating and extending the original study by Pollalis et al. we were able to gain a thorough understanding of users' visual behavior for the purpose of enhancing object-based learning. By adding eye tracking we found that users' visual behavior towards virtual learning material is similar to tangible ones. As mentioned above, users spent a significant amount of time looking at the manipulation features of the HoloLens interface. This highlights a need to further explore the role of physical manipulation in learning. We plan to extend this study to analyze collaborative object-based learning and how different interaction styles facilitate collaborative object-based learning.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation through grant OISE-1658594.

REFERENCES

[1] Ayman Alzayat, Mark Hancock, and Miguel Nacenta. 2014. Quantitative Measurement of Virtual vs. Physical Object Embodiment Through Kinesthetic Figural After Effects. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2903–2912. https://doi.org/10.1145/2556288.2557282

[2] Jorge Bacca, Silvia Baldiris, RamÃşn Fabregat, Sabine Graf, and Dr Kinshuk. 2014. Augmented Reality Trends in Education: A Systematic Review of Research and Applications. *Educational Technology and Society* 17 (10 2014), 133–149.

[3] Kuo-En Chang, Chia-Tzu Chang, Huei-Tse Hou, Yao-Ting Sung, Huei-Lin Chao, and Cheng-Ming Lee. 2014. Development and Behavioral Pattern Analysis of a Mobile Guide System with Augmented Reality for Painting Appreciation Instruction in an Art Museum. *Comput. Educ.* 71 (Feb. 2014), 185–197. https://doi.org/10.1016/j.compedu.2013.09.022

[4] Jill A. Cochran, Zane Cochran, Kendra Laney, and Mandi Dean. 2016. Expanding Geometry Understanding with 3D Printing. *Mathematics Teaching in the Middle School* 21, 9 (2016), 534–542. http://www.jstor.org/stable/10.5951/mathteacmiddscho.21.9.0534

[5] Eman Daraghmi, Chih-Tien Fan, Wei-Jun Lin, Zih-You Peng, and Shyan-Ming Yuan. 2015. Eye-Tracking Assisted Self-Directed Learning System. 52–54. https://doi.org/10.1109/ICALT.2015.33

[6] Christian Greiffenhagen and Stuart Reeves. 2013. Is Replication important for HCI? *CEUR Workshop Proceedings* 976.

[7] R. Gu, J. Zhang, J. Zhou, and M. Tong. 2014. The Baroque music's influence on learning efficiency based on the research of eye movement. In *2014 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*. 511–515. https://doi.org/10.1109/TALE.2014.7062619

[8] Martin Hachet, Benoit Bossavit, Aurélie Cohé, and Jean-Baptiste de la Rivière. 2011. Toucheo: Multitouch and Stereo Combined in a Seamless Workspace. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 587–592. https://doi.org/10.1145/2047196.2047273

[9] S. S. Hacisalihzade, L. W. Stark, and J. S. Allen. 1992. Visual perception and sequences of eye movement fixations: a stochastic modeling approach. *IEEE Transactions on Systems, Man, and Cybernetics* 22, 3 (May 1992), 474–481. https://doi.org/10.1109/21.155948

[10] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139 – 183. https://doi.org/10.1016/S0166-4115(08)62386-9

[11] Kasper Hornbæk, Søren S. Sander, Javier Andrés Bargas-Avila, and Jakob Grue Simonsen. 2014. Is Once Enough?: On the Extent and Content of Replications in Human-computer Interaction. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3523–3532. https://doi.org/10.1145/2556288.2557004

[12] David E. Irwin and Gregory J. Zelinsky. 2002. Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics* 64, 6 (01 Aug 2002), 882–895. https://doi.org/10.3758/BF03196793

[13] Karnyupha Jittivadhna, Pintip Ruenwongsa, and Bhinyo Panijpan. 2010. Beyond Textbook Illustrations: Hand-held Models of Ordered DNA and Protein Structures as 3D Supplements to Enhance Student Learning of Helical Biopolymers. *Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology* 38 (11 2010), 359–64. https://doi.org/10.1002/bmb.20427

[14] Larry Johnson, Samantha Adams Becker, Michele Cummins, Victoria Estrada, Alex Freeman, and Courtney Hall. 2016. *NMC Horizon Report: 2016 Higher Education Edition*. Technical Report. The New Media Consortium, Austin, Texas. https://www.learntechlib.org/p/171478

[15] Thomas Kador, Helen Chatterjee, and Leonie Hannan. 2017. *The materials of life: Making meaning through object-based learning in twenty-first century higher education*. UCL Press, 60–74. http://www.jstor.org/stable/j.ctt1xhr542.10

[16] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 1151–1160. https://doi.org/10.1145/2638728.2641695

[17] Andrew L. Kun. 2018. Human-Machine Interaction for Vehicles: Review and Outlook. *Foundations and TrendsÂő in Human-Computer Interaction* 11 (01 2018), 201–293. https://doi.org/10.1561/1100000069

[18] Andrew L Kun, Hidde van der Meulen, and Christian Janssen. 2017. Calling While Driving: An Initial Experiment with HoloLens. 200–206. https://doi.org/10.17077/drivingassessment.1636

[19] Christian Lander, Marco Speicher, Frederic Kerber, and Antonio Krüger. 2018. Towards Fixation Extraction in Corneal Imaging Based Eye Tracking Data. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. ACM, New York, NY, USA, Article LBW558, 6 pages. https://doi.org/10.1145/3170427.3188597

[20] Dorothea Lasky. 2009. Learning from Objects: A Future for 21st Century Urban Arts Education. *Penn GSE Perspectives on Urban Education* 6, 2 (2009), 72–76.

[21] Jiajia Li, Grace Ngai, Hong Va Leong, and Stephen CF Chan. 2016. Your eye tells how well you comprehend. In *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, Vol. 2. IEEE, 503–508.

[22] Paul Marshall. 2007. Do tangible interfaces enhance learning?. In *Proceedings of the 1st international conference on Tangible and embedded interaction*. ACM, 163–170.

[23] Stefan Mathe and Cristian Sminchisescu. 2015. Actions in the eye: Dynamic gaze datasets and learnt saliency models for visual recognition. *IEEE transactions on pattern analysis and machine intelligence* 37, 7 (2015), 1408–1424.

[24] Maria Montessori and Henry Wyman Holmes. 1912. *The Montessori Method: Scientific Pedagogy as Applied to Child Education in" The Children's Houses"*. Frederick A. Stokes Company.

[25] Susan M Munn, Leanne Stefano, and Jeff B Pelz. 2008. Fixation-identification in dynamic scenes: Comparing an automated algorithm to manual coding. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*. ACM, 33–42.

[26] Kenton P O'Hara and Stephen J Payne. 1998. The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive psychology* 35, 1 (1998), 34–70.

[27] Muzaffer Ozdemir, Cavus Sahin, Serdar Arcagok, and M Kaan Demir. 2018. The Effect of Augmented Reality Applications in the Learning Process: A Meta-Analysis Study. *Eurasian Journal of Educational Research* 74 (2018), 165–186.

[28] Samiullah Paracha, Toshiro Takahara, and Ayaka Inoue. 2016. Interactive screening of learning issues via eye-tracking technology. In *2016 13th Learning and Technology Conference (L&T)*. IEEE, 1–6.

[29] Christina Pollalis. 2017. HoloMuse Study. https://sketchfab.com/cpollali/collections/holomuse-study

[30] Christina Pollalis, Whitney Fahnbulleh, Jordan Tynes, and Orit Shaer. 2017. HoloMuse: Enhancing engagement with archaeological artifacts through gesture-based interaction with holograms. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, 565–570.

[31] Christina Pollalis, Elizabeth Joanna Minor, Lauren Westendorf, Whitney Fahnbulleh, Isabella Virgilio, Andrew L Kun, and Orit Shaer. 2018. Evaluating learning with tangible and virtual representations of archaeological artifacts. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, 626–637.

[32] Miguel A Recarte and Luis M Nunes. 2000. Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of experimental psychology: Applied* 6, 1 (2000), 31.

[33] P. Reuter, G. Riviere, N. Couture, N. Sorraing, L. Espinasse, and R. Vergnieux. 2007. ArcheoTUI - a Tangible User Interface for the Virtual Reassembly of Fractured Archeological Objects. In *Proceedings of the 8th International Conference on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST'07)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 15–22. https://doi.org/10.2312/VAST/VAST07/015-022

[34] Metod Rybar and Maria Bielikova. 2016. Automated detection of user deception in on-line questionnaires with focus on eye tracking use. In *2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*. IEEE, 24–28.

[35] Dario D Salvucci and Joseph H Goldberg. 2000. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*. ACM, 71–78.

[36] Kamran Sedig, Maria Klawe, and Marvin Westrom. 2001. Role of interface manipulation style and scaffolding on cognition and concept learning in learnware. *ACM Transactions on Computer-Human Interaction (TOCHI)* 8, 1 (2001), 34–59.

[37] Ram Singh, Sushil Chandra, Kalyani Dhusia, and Greeshma Sharma. 2016. Capacitating surveillance and situational awareness with measure of visual engagement using eyetracker. In *2016 International Conference on Computing, Communication and Automation (ICCCA)*. IEEE, 1161–1165.

[38] Gunnvald B Svendsen. 1991. The influence of interface style on problem solving. *International Journal of Man-Machine Studies* 35, 3 (1991), 379–397.

[39] Hidde van der Meulen, Andrew L Kun, and Orit Shaer. 2017. What Are We Missing?: Adding Eye-Tracking to the HoloLens to Improve Gaze Estimation Accuracy. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*. ACM, 396–400.

[40] Hidde van der Meulen, Petra Varsanyi, Lauren Westendorf, Andrew L Kun, and Orit Shaer. 2016. Towards understanding collaboration around interactive surfaces: Exploring joint visual attention. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 219–220.

[41] Peter Vorderer, Werner Wirth, Feliz R Gouveia, Frank Biocca, Timo Saari, Futz Jäncke, Saskia Böcking, Holger Schramm, Andre Gysbers, Tilo Hartmann, et al. 2004. MEC spatial presence questionnaire (MEC-SPQ): Short documentation and instructions for application. *Report to the European community, project presence: MEC (IST-2001-37661)* 3 (2004).

[42] Esra Yecan and Kursat Cagiltay. 2006. Cognitive styles and students' interaction with an instructional web-site: Tracing users through eye-gaze. In *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*. IEEE, 340–342.