Do ArUco Markers Pose a Distraction?

Kate Baringer
Psychology Department
Clemson University
Clemson, SC
kmbarin@clemson.edu

Matthew Beach
Department of Computer Science
Clemson University
Clemson, SC
mbeach@clemson.edu

Joshua Lacy
Department of Computer Science
Clemson University
Clemson, SC
jrlacy@clemson.edu

ABSTRACT

ArUco markers are a useful tool in 3D eye tracking. They allow the computer to calculate gaze location within the space. However, these markers may have an adverse effect on participants as they may unintentionally attract their gaze. We conducted this experiment to test whether, while observing and memorizing details about an art piece, participants spent time looking at the ArUco marker. The hypothesized outcome was that the participants would initially spend a small amount of time on them in the early stimuli but less visual attention would be given to the markers on the later stimuli.

KEYWORDS

Art, ArUco, Eye tracking, Cueing, Gazepoint

1 Introduction

In 3D eye tracking, ArUco markers are often used to streamline the process of estimating marker poses with six degrees of freedom [1]. These QR-code-like black and white square markers act as anchoring points for the eye tracking camera, allowing eye tracking software to calculate where in 3D space gaze in being allocated relative to these markers. Often used to define areas of interest (AOIs), ArUco markers are an important tool used during 3D space eye tracking data analysis. However, it is possible that the pattern of where these markers are placed within an experimental environment may act as cues for participants, drawing the eye to different AOIs because of the markers' presence rather than because the participant believes information relevant to their task can be found there.

1.1 Background

When discussing gaze and visual attention, it is important to first recognize that something cannot be the object of visual attention if it is not first looked at. Alternatively, just because something was looked at does not mean that it was attended. For example, the looking-but-not-seeing phenomenon, in which humans fail to notice a change in their environment even though it was looked at because it wasn't relevant to the task they were trying to accomplish [2]. This is a form of attentional blindness, which can come in the form of either ignoring constant stimuli in the environment or failing to notice non-task-relevant changes [3]. It has been shown that both the familiarity and meaningfulness of stimuli affect whether or not it is perceived, with unfamiliar meaningless stimuli being unlikely to be perceived even if it was looked at [4]. For participants unfamiliar with what ArUco markers are and how they're used, which is likely to be most participants, we believe these markers will fall under the "unfamiliar and meaningless" category.

When performing visual search, several factors influence gaze patterns. There are bottom-up factors like saliency, how much an object stands out from its surroundings, but there are also top-down processes where experience with either that particular scene or similar scenes will inform search strategies [5]. One of these top-down processes is call contextual cueing. In contextual cueing, repeated features of the target stimuli, for example always appearing in a similar position within the search field or consistently being accompanied by another type of object, called co-location, will help facilitate the search and reduce the time it takes to find the target stimuli [6]. For example, if you're looking for a computer mouse you may know that they are often located near a computer, so when walking into a new office space you could use this knowledge to quickly find the computer mouse by first finding the larger and easier to see co-located object – the computer. While this type of background knowledge, or schema, is generally useful, it can also cause unintended learned cueing within experimental or training designs [7]. This kind of unintentional cueing, due to the necessary co-location of ArUco markers with stimuli-of-interest within a 3D environment, is the potential concern this paper hopes to address.

1.2 Hypothesis

The purpose of this paper is to determine whether ArUco markers draw the gaze or whether they are inconspicuous enough to "fade into the background". In this study, the location of the ArUco marker will be consistent relative to the stimuli, and we will not be explaining what the marker is or why it is there. Due to this combination of location consistency and the ArUco marker's lack of meaning to the participant, we hypothesize that, while the ArUco markers may draw the gaze during the first few sets of stimuli, it will eventually "fade into the background" and be ignored by participants. This is based on the previously discussed trend for objects in the environment not to be perceived when they

are unfamiliar or meaningless to the observer and irrelevant to the task being accomplished [3]. We anticipate that this trend will outweigh the potential for a schema to be developed based on the co-location of the ArUco marker and task-relevant information.

2 Methods

2.1 Participants

A group of 20 participants, 10 male, 9 female, 1 non-binary took part in our study. Their ages ranged from 19 to 45 with the average being 25. 27.78% of the participants had previous experience with ArUco markers. Two participants were excluded from the analysis due to their moving out of range of the eye tracker during the experiment, leaving us with 18 participants, 9 male, 8 female, 1 non-binary.

2.2 Stimulus

Three different forms of stimuli were used. The first stimulus used was a picture of a painting and the description of the painting merged, as shown in Figure 1. The second stimulus, seen in Figure 2, used the merged image of the painting and description with a small ArUco marker above the description. The third stimulus, Figure 3, used the merged image with a large ArUco marker above the description.

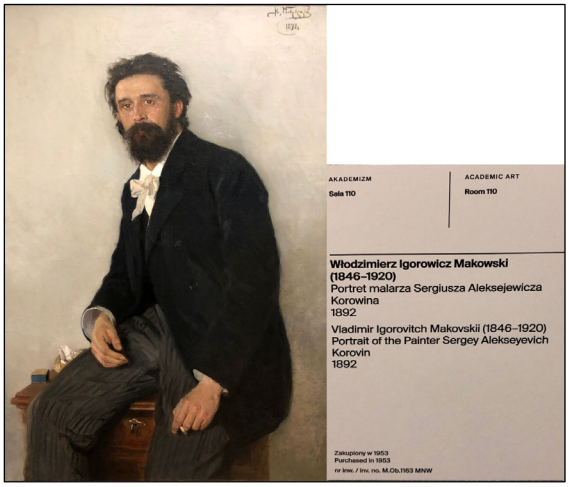

Figure 1: Example of a stimulus used for the "absent ArUco marker" condition.

Figure 2: Example of a stimulus used for the "small ArUco marker" condition.

Figure 3: Example of a stimulus used in the "large ArUco marker" condition.

2.3 Apparatus

The used apparatus is a Gazepoint GP3 eye tracker. It has 1% accurate and a 60 Hz sampling rate. Participants sat approximately 30 inches from the eye tracker.

2.4 Experimental Design

The experiment utilized a factorial between-subjects design with one independent variable and three conditions. The independent variable was the ArUco marker presence with levels “absent”, “small”, and “large”. Participants were randomly assigned to one of the three groups such that the ArUco marker level was constant for his/her trial. This way, the participants would not look at the ArUco markers simply because they varied in size or presence throughout the trial. All participants were given the exact same sets of images and the same visual search task. The order in which the paintings were presented was randomized. The dependent

variables included number of fixations on the ArUco marker, percent dwell time on the ArUco marker, and maximum duration of fixations on the ArUco marker for each stimuli presented. A total of ten sets of painting, description, and ArUco marker (or lack of) were presented to each participant during the experiment.

2.5 Procedure

Participants were greeted by the researchers and taken to the computer lab where the experiment was conducted. The participants read the informational letter and filled out the pre-questionnaire which asked age, gender, use of corrective lenses, and pre-existing eye conditions. After filling out the forms, the participant could ask any questions before beginning the experiment.

The experiment began with calibration of the eye tracker using Gazepoint. After successful calibration, an instructional page popped onto screen to inform the participant of their visual search task. They were told that they would be given 30 seconds to look at each painting, noting the author and title, then they would be asked to describe what they saw to the experimenter in 30 seconds or less. Participants were given a chance to ask any questions after reading the instructions. Next, the participant was presented with a series of 10 stimuli, one at a time, and after each stimulus they were given another 30 seconds to describe the painting in their own words to the experimenter.

After concluding the experiment, the participants were given a post-study survey asking what they thought the purpose of the study was and whether they had any experience with either eye tracking or ArUco markers before. Next, they were debriefed, told the true purpose of the study, thanked for their contributions, and dismissed from the experiment. In total, the experiment lasted about 15 minutes.

RESULTS

All statistical analysis was done using SPSS version 27 with a significance level of $\alpha = 0.05$. Figure 4 shows an example of how areas of interest (AOIs) were set up in Gazepoint for analysis. Multivariate analysis was done using AOI Name (Image, Description, Marker), and Condition (Small, Medium, Large) as independent variables. Dependent variables were the percentage of time each AOI was viewed, the number of seconds each AOI was viewed, and the number of fixations on each AOI.

Figure 4: Example of AOIs used in the "large ArUco marker" condition. Blue is AOI "image", green is AOI "marker", red is AOI "description".

The overall effect of AOI name was significant, Pillai's Trace $\eta = 0.844$, $F(2) = 1429.28$, $p < 0.001$. The overall effect of condition was not significant, however there was a significant interaction between AOI and condition, $\eta = 0.157$, $F(8) = 11.29$, $p < 0.001$. The non-significant between-conditions differences as well as the interaction can be seen in Figures 5 and 6, which show the breakdown of the number of fixations for AOI and the percent of time viewed for each AOI by condition.

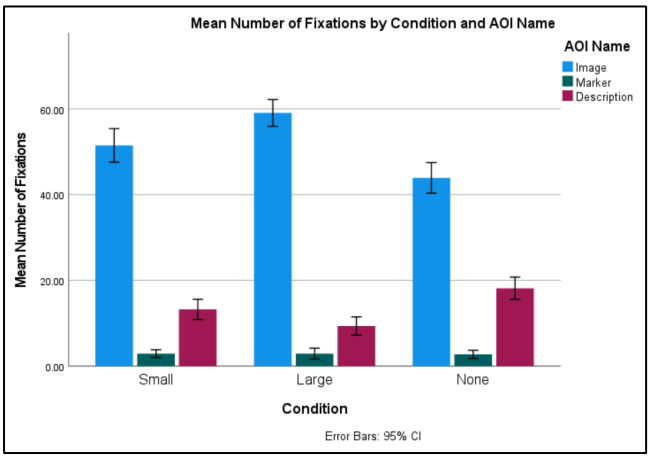

Figure 5: Number of fixations per AOI across conditions

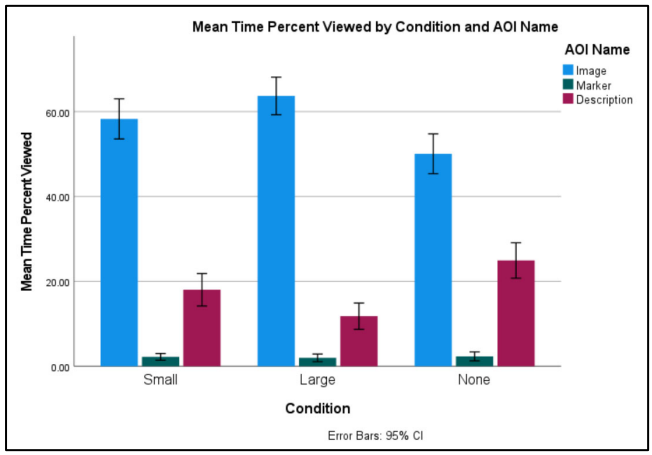

Figure 6: Percent each AOI was viewed by condition and AOI type

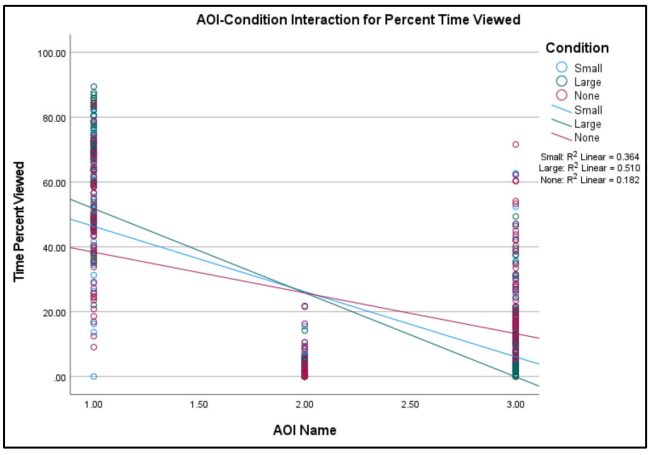

Figure 8: Interaction between AOI and Condition for Percent Time Viewed

Percent Dwell Time

Between-subjects effects of percent dwell time showed that percent dwell time was significantly different between AOI names $F(2) = 804.98$, $p < 0.001$. The interaction between AOI name and condition was also statistically significant for percent dwell time $F(4) = 15.014$, $p < 0.001$.

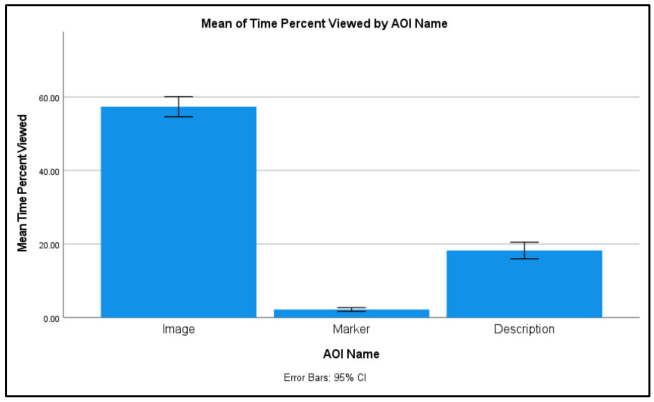

Figure 7: Mean percent of time each AOI was viewed by AOI

Seconds Viewed

Between-subjects effects of seconds viewed showed that the number of seconds each AOI was viewed differed significantly by AOI name $F(2) = 804.97$, $p < 0.001$. The interaction between AOI and condition was also statistically significant for the number of seconds viewed $F(4) = 15.014$, $p < 0.001$.

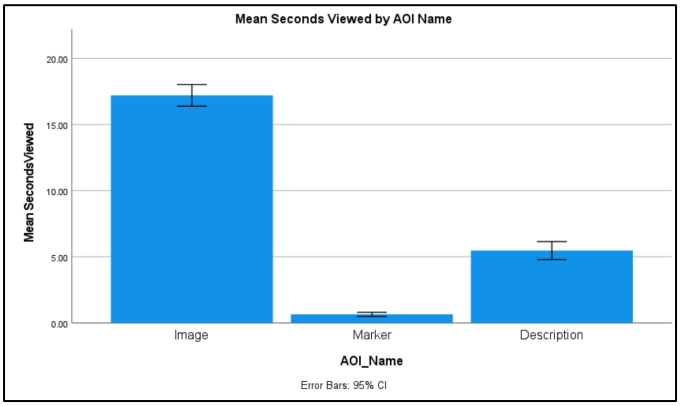

Figure 9: Mean seconds each AOI was viewed by AOI

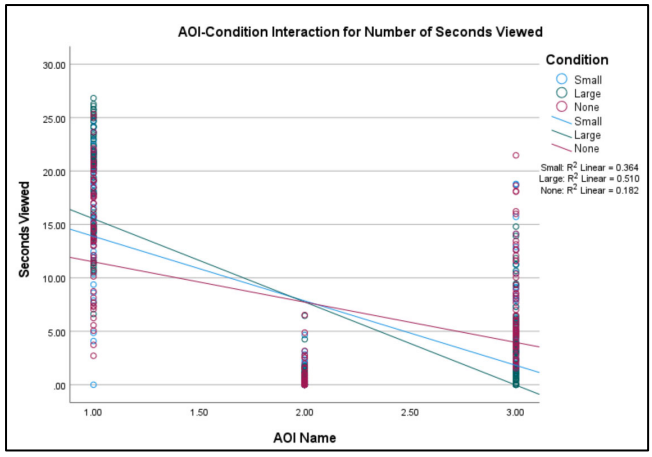

Figure 10: Interaction between AOI and Condition for number of seconds each AOI was viewed

Number of Fixations

Between-subjects effects of the number of fixations per AOI showed that the number of time each AOI was looked at did differ significantly based on AOI name $F(2) = 1211.74$, $p < 0.001$. The interaction between AOI and condition was also significant for number of fixations $F(4) = 22.7$ $p < 0.001$.

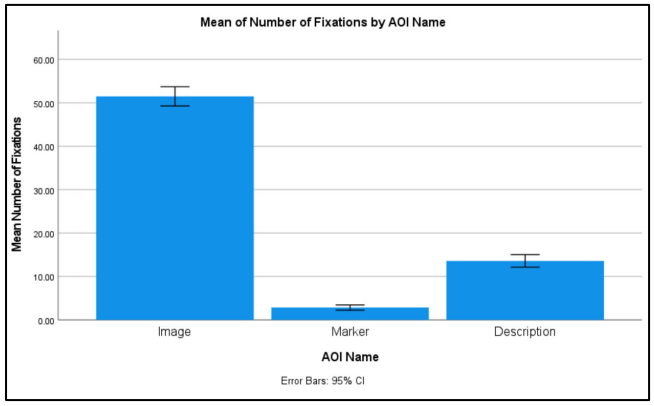

Figure 11: Mean number of fixations per AOI by AOI

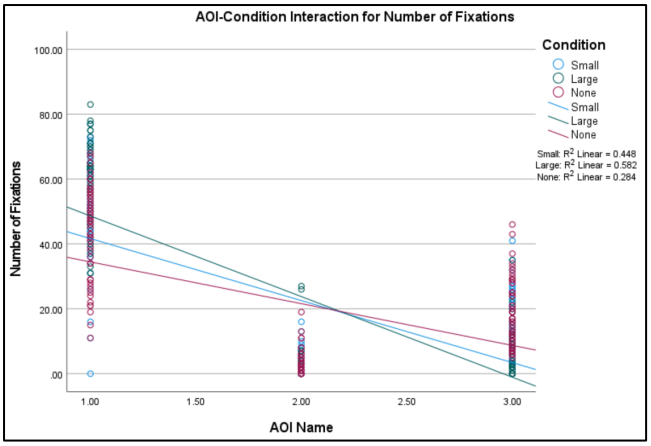

Figure 12: AOI-Condition interaction for number of fixations per AOI

Noticing

While the effect of condition was not significant on how often the ArUco markers were looked at, there were interesting trends seen in noticing of the markers. When asked whether they remembered looking the ArUco markers during the experiment, three respondents in the small maker group said "never", one said "rarely, and two said "sometimes. In the large marker group, however, only one participant said "never", three said "rarely", one said "sometimes, and one said "often".

DISCUSSION

Before the data was run through the statistical analysis, each participant's data was analyzed for quality. After reviewing the data, two participants' data was excluded from the study due to falling out of frame of the eye tracker, resulting in missing data for those participants.

The given task of "viewing the image in order to describe it, noting the artist and title" had a large effect on the resulting data. Due to the task and amount of information contained in the image to complete the task, the mean of all participants in all conditions spent a plurality of the time (> 45%) looking at the image in each stimulus, as shown by Figure 6. After the image, the description was the next most viewed, as it contained useful information such as the name of the painting, the artist's name, and the year created. However, this data was usually only viewed once and memorized, not needing to be viewed again. The least viewed was the ArUco markers, or empty space in the condition that had no markers. Viewing some of the results manually, oftentimes the marker was only viewed as a stop between the image and the description. Going from the top of the image to the description in the bottom right corner of the stimulus would travel through the area of the ArUco marker, resulting in more fixations and view time for that AOI.

The condition level of the ArUco marker was not significant. The null hypothesis that ArUco markers do not affect the gaze pattern (number of fixations, percent time viewed, and total time viewed)

for the task cannot be disproven based on the data collected. Therefore it's reasonable to say that the presence of an ArUco marker does not distract the participant from the task at hand. Figure 5 shows the consistency between conditions and how much the marker AOI was viewed. Occasionally, the marker AOI was never viewed during the experiment.

These results suggest that the presence of ArUco markers is not significantly disruptive to the experimental task when they are present, which is promising news for 3D eye tracking studies that require these to be present, but the experimenters may be concerned that they are distracting to the participant. While the size of the ArUco marker may make a difference in noticeability, as demonstrated by the increased number of participants that remembered seeing ArUco markers in the large marker condition compared to the small marker condition, the actual number of fixations on the ArUco marker as well as the percent of time spent looking at the ArUco marker was not significantly different between the two conditions. This again has promising implications for 3D eye tracking studies.

We recommend that this study be duplicated with a larger sample size to confirm the findings. While our findings were significant, we still believe that these findings would be better supported if replicated in another, larger study. Future research recommendations include determining whether ArUco markers may act as cues for participants.

CONCLUSION

The experiment conducted provides evidence that the presence of ArUco markers do not have a significant effect on an individual's search with the given task. This suggests that studies using ArUco markers aren't seriously impacted by using them and is a safe element to use in other visual search experiments. The experiment however can't be considered conclusive because of a low sample size, the use of specific directions, and the ArUco markers being distinctly separate from any other on-screen element. This study also does not provide any evidence as to whether ArUco markers could act as cues within a visual search area.

ACKNOWLEDGMENTS

We would like to recognize Professor Andrew Duchowski for his guidance throughout the process of conducting this research.

REFERENCES

[1] S. Garrido-Jurado, R. Munoz-Salinas, F.J. Madrid-Cuevas and M.J. MartinJimenez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. In *Journal Pattern Recognition Vol 47, Issue 6,* (June 2014), 2280-2292. https://doi.org/10.1016/j.patcog.2014.01.005.

[2] Kirsten Cater, Alan Chalamers and Patrick Ledda. 2002. Selective quality rendering by exploiting human attentional blindness: looking but not seeing. In *Proceedings of the ACM symposium on Virtual reality software and technology (VRST '02)*. ACM Press, New York, NY, 17-24. DOI: https://doi.org/10.1145/585740.585744.

[3] Daniel Simons. 2000. Attentional capture and inattentional blindness. In *Trends in Cognitive Sciences (Vol 4 Issue 4)*.DOI: https://doi.org/10.1016/S1364-6613(00)01455-8.

[4] Arien Mack, Zissis Pappas, Michael Silverman and Robin Gay. 2002. What we see: Inattention and the capture of attention by meaning. In *Journal of Consciousness and Cognition (Vol11 Issue 4).* 488-506. DOI: https://doi.org/10.1016/S1053-8100(02)00028-4.

[5] J.M Wolfe, S.J. Butcher, C. Lee and M. Hyle. 2003. Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance (29)*. 483-502. DOI: https://doi.org/10.1037/0096-1523.29.2.483.

[6] Marvin Chun. 2000. Contextual cueing of visual attention. In *Trends in Cognitive Sciences (Vol 4 Issue 5)*. 170-178. DOI: https://doi.org/10.1016/S1364-6613(00)01476-5.

[7] Adam Biggs, Dominick Pistone, Mackenzie Riggenback, Joseph Hamilton and Kara Blacker. 2021. How unintentional cues can bias threat assessments during shoot/don't-shoot simulations. In *Journal of Applied Ergonomics (Vol 95)*. DOI: https://doi.org/10.1016/j.apergo.2021.103451.