GUIDED PROJECT IN BIOLOGY RESEARCH – HONOR'S PROGRAM – 72328 OCTOBER 2014
THE HEBREW UNIVERSITY

SEMANTIC AND VISUAL ENCODING IN INCIDENTAL MEMORY FOR OBJECTS DURING VISUAL SEARCH IN NATURAL SCENES

Gal Nissim
Advisor: Prof. Ehud Zohary

ABSTRACT

Humans must be able to efficiently find objects in complex environments in order to act in the world. The aim of this research is to find out whether object-location binding in memory is formed for 'target objects', which were incidentally gazed upon while searching for other 'search objects'. To that end, we examined incidental memory for objects and for their position in familiar natural scenes (i.e. living room). Participants performed a visual search task, followed by a memory test, which included a recognition test and a position task for 'target objects'. The results indicate that even though participants were not instructed to search for these target objects, information about them was yet retained after the search (semantic and visual representation, as well as location). These results demonstrate binding to location in visual and semantic memory that was acquired without intention.

INTRODUCTION

One of the most common behaviors in which humans engage is visual search (Hollingworth, 2012). From the moment we open our eyes in the morning we start to search. One can look for a cup in order to make coffee, for keys before leaving the house, or for any other needed object during the day. The ability to find objects in our environment is essential for our daily conduct. Searches are repeated many times for common tasks in environments we inhabit, which remain predictable despite their variability. This raises the possibility that search through scenes is strongly influenced by learning (Chun, 2000; Võ & Wolfe, 2013).

Humans rely on their visual perception continuously (Ricciardi et al., 2009). Vision allows us to perceive the shape of objects, to classify them into meaningful categories and to use them properly. Objects from certain categories are favored to be found at appropriate locations. Hence, we can accurately perceive the spatial location of objects (Hollingworth, 2005). For example, pots tend to be found in kitchens, so a search for a pot will be conducted much differently if one is searching for the pot in a kitchen as opposed to a living room (Henderson et al., 1999; Neider & Zelinsky, 2006; Torralba et al., 2006). Furthermore, while searching for a pot in a kitchen, one will prefer to search for it on the stove-top rather than the kitchen floor. This indicates that we use prior knowledge while searching. When searching for an object that one has never seen before, he or she would use prior knowledge of the object's template. Regarding the pot example, one will search for a round shiny object with two handles. That representation of the target object must be retrieved from one's memory (Yang & Zelinsky, 2009).

While searching for an object, one might try to retrieve a vague memory of the last place it has been seen (Võ & Wolfe, 2013). Memory for objects in natural scenes has been widely investigated. It was found that people can remember the specific context of the scene as well as the position of objects in the same scene (Hollingworth, 2005).

It seems intuitive to think that previous exposure or interaction with an environment should make it easier to search through it. However, a recent study conducted by Võ & Wolfe demonstrated that repeatedly searching for multiple different objects in the same unchanging scene does not dramatically speed up the search, despite the observer’s increasing familiarity with the scene and gazing on distracters incidentally. Only when observers were asked to search for the same object again did the search become considerably faster (Võ & Wolfe, 2012; Võ & Wolfe, 2013).

This lack of finding a behavioral benefit of incidental fixation seems odd, since previous studies have shown that incidental distractor fixations during search increase recognition memory for them. Namely, incidental memory does exist (Castelhano & Henderson, 2005; Võ et al., 2008). We tend to encode information from a particular scene, such as the spatial structure of the environment and the locations of individual objects. We are capable of remembering these features of the environment (Hollingworth & Henderson 2002).

According to Paivio, both visual and verbal information are used to represent information (Sternberg, 2003). The mental codes corresponding to these representations are used to organize incoming information that can be acted upon, stored, and retrieved for subsequent use. One can retrieve either the word or the image separately, or both simultaneously (Paivio, 1991). Therefore, we would expect to have differences in the subjects' performance depending on the type of encoding.

In this paper we seek to test the formation of object-location binding in incidental memory during visual search (in natural scenes). We argue that familiarity with a scene or previous encounters with objects embedded in it will form memory for those objects and their location. Therefore, we examine how visual encoding, semantic encoding and eye-movement while searching affect the binding of memory for position of objects and their identity.

METHODS

PARTICIPANTS

12 subjects participated in the experiment (ages 18-31, Mean Age = 24.53; 6 females and 6 males). All subjects were naive and had normal or corrected to normal visual acuity by self-report and normal color vision as assessed by the Ishihara test. All subjects were paid volunteers who had given written informed consent. Experimental procedures were approved by the ethics committee of the Hebrew University of Jerusalem. The subjects were divided into two experiment groups.

STIMULI

4 full-color images of indoor scenes were presented in the sessions of the experiment. An additional image was used for practice trials. The scenes were formed in Google SketchUp, a 3D modeling program for applications. Each scene included 20 different 'search objects' [1] and 10 singleton 'target objects' [2] '. Only one representation of a specific 'search object' was presented in the same scene. In addition, there were no repetitions of the 'target objects' between the scenes.

For both 'search objects' and 'target objects', different feature variables were chosen so that some of the objects were more salient than others (color, size, location

[1] 'Search objects' were the objects participants searched for during the search task.

[2] 'Target objects' were present in the scenes in which the search task was conducted, but the subjects were not asked to search for them.

and scene context). That way we created varying levels of difficulty in the search task, and variance in memory for objects.

Scenes were displayed on a 20 inch computer screen (resolution 1280x1024 pixel, 85 Hz) subtending visual angles of 360º (horizontal) and 270º (vertical) at a viewing distance of 55cm (H: 1º = 35.56 pixels; V: 1º = 37.93 pixels). The experiment was conducted in a dimly lit room, subjects sat in front of the computer screen while their head was held in a chinrest.

Figure 1. A living room scene in which one is required to look for a bottle of water, a rubber duck, and so forth. The scene includes two kinds of objects: ‘search objects’ (i.e. bottle of water) and ‘target objects’ (i.e. a dog, which are not actively searched for during the "search" stage).

EYE-TRACKING

Eye movements were recorded with a video-based, infra-red, desk-mounted eye tracker (Eye Link1000, SR Research, Ontario, Canada) at a sampling rate of 1000 Hz. Viewing was binocular but only the position of one eye was tracked. The manufacturer's software was used for stimuli presentation, calibration, validation, drift correction, and determination of periods of fixation. Eye-position data was used to automatically monitor online performance of the task. Eye-movement data was analyzed using the manufacturer's software (Experiment Builder, SR, Research, Ontario, Canada) SPSS and Matlab R2013b.

EXPERIMENTAL DESIGN

The experiment included two parts: a search task and memory test. Before each scene a randomized 9-point calibration and validation procedure was performed. Before each new 'search object', a drift correction was applied and if necessary, a re-calibration and re-validation procedure was performed.

The experiment included 4 scenes in total. Each scene contained two kinds of objects: 'search objects' and 'target objects' (i.e. distractors). The scenes were divided between the two experiment groups so that each group had searched in only half of the scenes, but had a memory test on 'target objects' from all of the scenes. In that way novel 'target objects' for one group were the 'target objects' for the other group (and vice versa). Thus, the groups were counterbalanced for each 'target object' and that enabled to compare between its two conditions (appeared and novel).

SEARCH TASK:

In the search task participants searched, one after another, for 20 different 'search objects' in the same unchanging scene (order of searches was randomized). The 'search object' was identified at the start of each search by presenting it textually (font: Times New Roman, font size: 48, color: black, duration: determined by a button click) in the center of a white screen. Participants were instructed to search for the object, and once found, to press a button while fixating on the object as fast as possible. Transition to the next search target was enabled only if the subject was fixating on the correct 'search object'. The next search started with the appearance of a new word on a white screen after a drift correction was made (see Figure 2).

After 20 searches in the same scene, a calibration and validation procedure was performed before starting the search in the next scene (order of scenes was randomized). After searching through two different scenes (the first part of the experiment), a memory test was followed. Participants were not aware that they would have to go through a memory test.

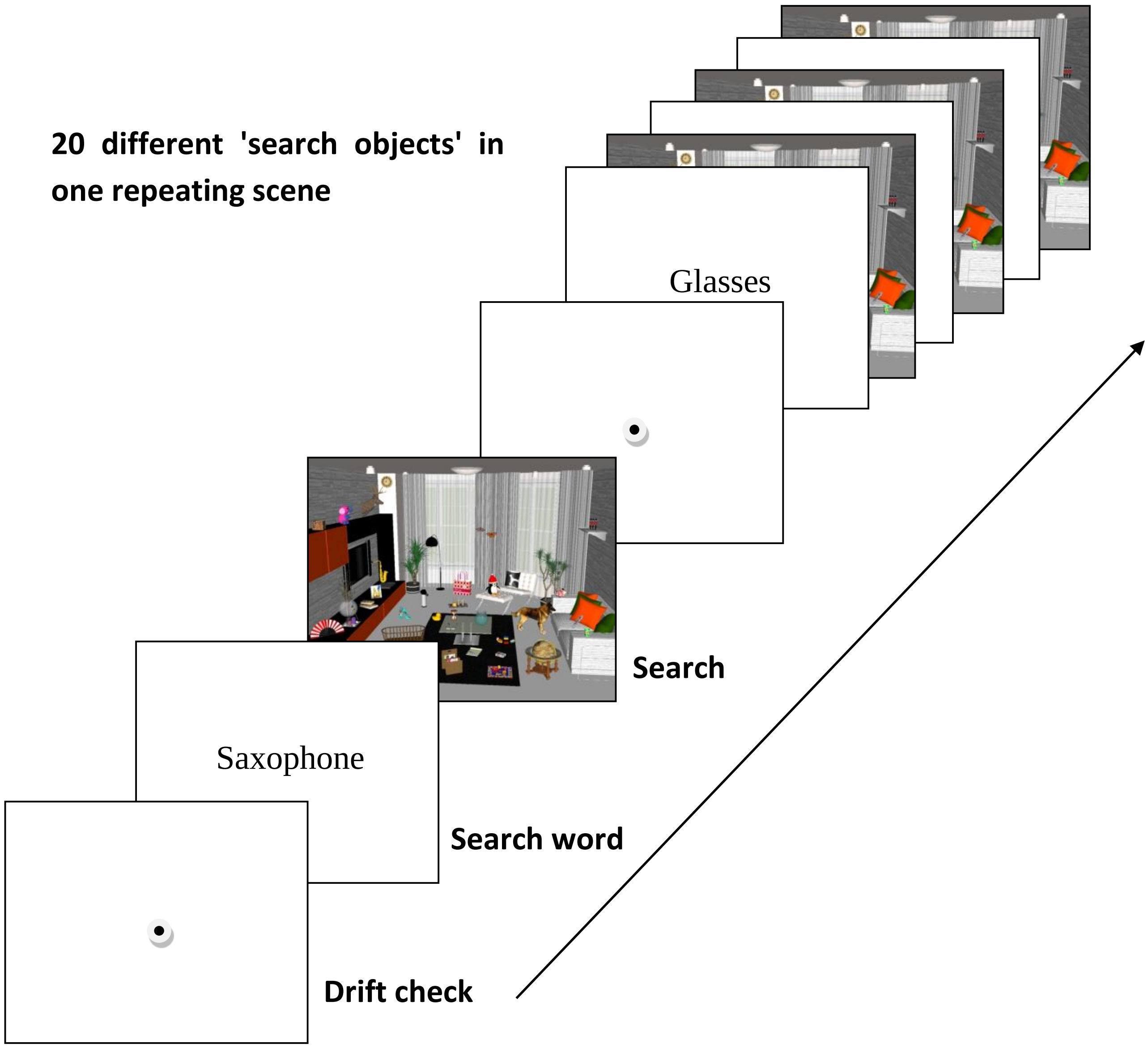

Figure 2. Search task: 20 repeated searches in the same scene. Search words at the center of a white screen designated the object of search for each search.

MEMORY TEST:

The second part of the experiment, the memory test, was divided into two tasks: recognition task and position task. Each memory test started by presenting a semantic representation of an object - a target word (font: Times New Roman, font size: 48, color: black, duration: determined by a button click). The participants were asked to determine whether that object already appeared in one of the former scenes or whether it is a novel object. The decision was made by pressing a button (a binary yes/no response). See Figure 3: 1.a & 2.a.

As was mentioned before, the 'target objects' that did appear were present in the scenes in which the search task was conducted, but the subjects were not asked to search for them.

If the object did appear (Figure 3: 1.a-d), a white screen with four figures from the same category around the screen center was displayed. In addition, feedback and instructions were presented. An image of the correct object was displayed in its original size (the other distractor figures were at the same size and point of view). The participants were asked to select the correct object that appeared amongst the four items by pressing the mouse. This was defined as a visual representation of a ‘target object’.

In that way, the participants were tested both on semantic and visual representations for each ‘target object’ that appeared (see Figure 3: 1.a-b). After that, the figure of the correct item was presented in the screen center, while a feedback and instructions were presented too. When the subjects pressed a button it would lead them to the position task.

In the position task an ‘empty scene’ was displayed. An ‘empty scene’ is the original room of the 'target object' without the objects in the room (see Figure 3: 1.d). The subjects were asked to position the 'target object' in the ‘empty scene’ (using the mouse), based on their memory of its original location in its original scene.

If the object did not appear, after a feedback was given, a figure of the object would be presented in the center of a white screen. When the subject pressed a button it would lead to the position task, and an ‘empty scene’ would be displayed (Figure 3: 2.a-d). In this case the ‘empty scene’ was not one of the previous scenes that appeared in the search task. The participants were instructed to position the novel object based on their own judgment in the new ‘empty scene’ (see Figure 3: 2.d). All feedbacks, instructions and semantic representations of an object were displayed inside a grey textbox (size: 750X200 pixel) at the top of a white screen.

In total, the participants were tested on 40 'target objects', 20 of them appeared in one of the former scenes (10 for each scene) and the other 20 were novel (order of objects was randomized). At the beginning of the experiment, as part of the search task, a practice trial was made by the subjects. This learning trial was not included in the final analyses. Each experiment lasted approximately 30 minutes.

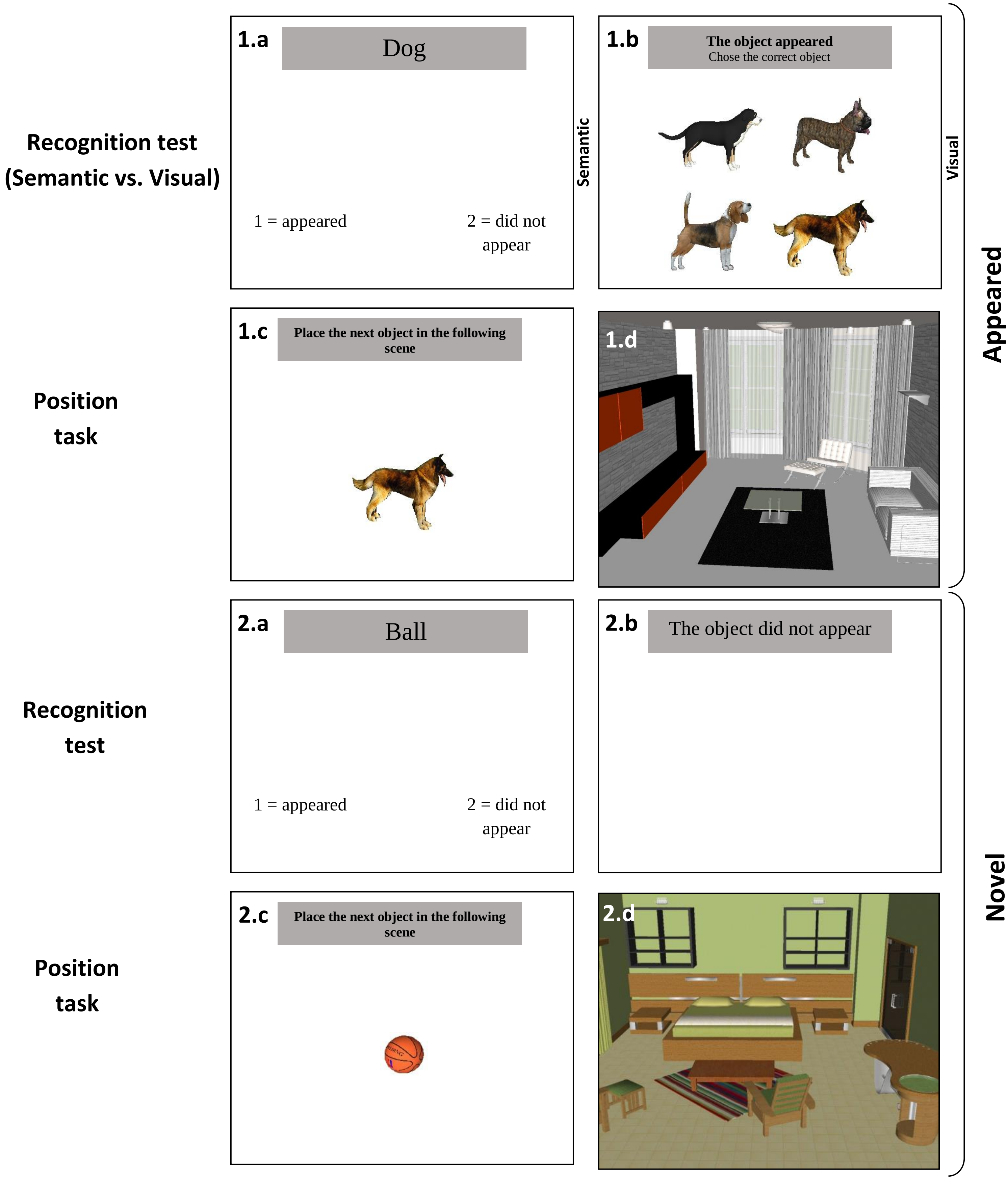

Figure 3. Memory test: 40 objects were included in the memory test. Half of them appeared before and half of them were novel. No. 1 describes an object that appeared before and No. 2 describes a novel object. The memory test is divided into two parts: recognition test and position task. For the appeared objects the recognition test contained two kinds of representations (semantic and visual). Subjects were asked to place the objects in the 'empty scenes' (1.d &2.d) in the place that they appeared or are likely to be (if they did not appear.

DATA ANALYSIS

Raw data was filtered using SR Research Data Viewer (SR Research, Ontario, Canada), and all analysis was done using Matlab in house code (The MathWorks, Natick, MA) and Microsoft Excel (Microsoft, Redmond, WA). Statistical analysis was done using SPSS (version 21.0 for Windows, IBM, Armonk, NY).

EYE-TRACKING ANALYSIS

The small interest area for each object, of both kinds (‘search objects’ and ‘target objects’), was defined by a rectangular box that was large enough to include that object. A minimum threshold for interest areas was defined (1.5X1.5 degrees). If limits of an object’s interest area were under that threshold, the limits' size were corrected (1.5º (horizontal) = 53.34 pixels, 1.5º (vertical) = 56.895 pixels). All values were rounded off.

The large interest area for each 'target object' was defined by adding to its small interest area half a degree in each direction (0.5º (horizontal) = 17.78 pixels, 0.5º (vertical) = 18.965 pixels). All values were rounded off.

Because of the accuracy of the EyeLink 1000 (Head Supported: 0.25º-0.5º average accuracy (EyeLink 1000 Technical specifications, 2013) the large interest areas were determined to be the sufficient areas for fixation, in order to move to the next search. The small interest areas are more accurate and therefore they were used in the analysis.

In order to investigate search performance, eye movement behavior and memory, a set of measures was calculated: *response time* was defined as the time that elapsed from the scene display onset (after the target word screen offset), until the button click while fixation was made on the correct object. *Incidental gaze* time was defined as the amount of time spent on looking at an object while it was not the target. We measured Incidental gaze time for the two types of objects. Incidental gaze time for the ‘search objects’ was calculated by summing up the time spent fixating on an object’s interest area, until that item became the object of search. So fixations on that ‘search object’ that followed its search trail were excluded. Incidental gaze time for the ‘target objects’ was calculated by summing up the amount of time spent fixating on an object’s interest area during all the trails of the ‘search task’ in its original scene.

DATA EXCLUSION

Errors could be made by looking at a location outside the target interest area, while pressing the button or by failing to find the target of search after an extended period of time (RT > 55,000 msec). Those trails were defined as error trails (5% of the experiment data).

Data of response time and incidental gaze time was analyzed using IBM SPSS Statistics and Matlab In order to identify outliers for each subject. The values that were larger than 4 standard deviations above the mean were defined as outliers and were excluded (2.2% of the experiment data).

3 of the 15 subjects were excluded from the analysis: 2 of them due to lack of understanding the instructions and moving a lot during the experiment. One of them, due to significantly long response time (RT > 250,000 msec) in one of the trials of the 'search task'. The long response time was a result of carrying out a serial search by the subject after she did not manage to find a search target.

BEHAVIORAL ANALYSIS

In order to investigate memory performances the 'target objects' were divided into a set of measures (hits, misses, correct-recognitions, false-alarms) and signal detection theory estimates of discriminability (d') and response bias (beta) were calculated for the semantic objects representation. For the visual representation hits and misses were calculated (for semantic vs. visual see Figure 3: 1.a-b).

The position error was defined as the Euclidian distance between the center of a 'target object' and its reposition by a subject. The position error was calculated in pixels.

RESULTS

INCIDENTAL GAZE TIME AND RESPONSE TIME

If some memory is used in the search task, large incidental gaze time should reduce the response time of the search. In order to examine that, we plotted the RTs of every 'search object' by every observer against the sum of all fixation durations on each object prior to it becoming the search target (i.e. incidental gaze time). 4 data points with gaze duration larger than 4 standard deviations above the mean of each subject were excluded. The results are seen in Figure 4. Incidental gaze time does not reduce search time.

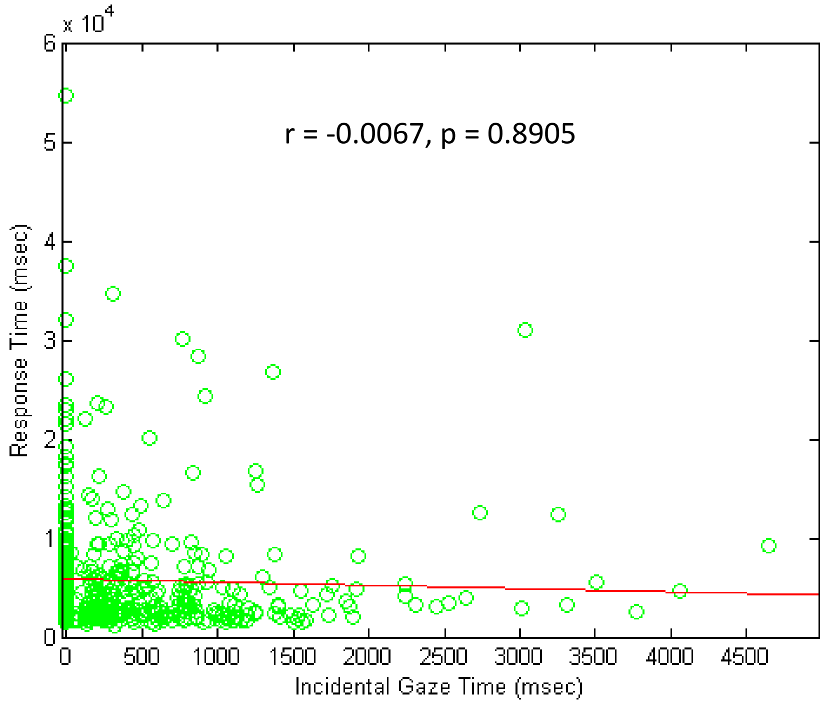

Figure 4. Previous incidental exposure to objects does not improve search performance. Search time for all stimuli across all participants as a function of the amount of time that a target of search was incidentally fixated upon in previous trials.

It seems counterintuitive that search time does not decrease as time of former exposure to the object increases and accumulates during the searches through the same scene. Previous research supports this finding (see Võ & Wolfe, 2012).

However, beside the formation of memory for the objects' identity during the search (e.g., Castelhano & Henderson, 2005; Võ et al., 2008), we wished to examine memory for their locations. If there exists an object-location binding in the formation of incidental memory, perhaps previous exposure time would affect the latter, without necessarily modifying search response time. In order to examine that, we used several tests.

POSITION ERROR AND MEMORY

Each 'target object' had two conditions in which participants located that object: appeared and novel. The appeared condition refers to the case in which the 'target object' was present in one of the previous scenes in the search task, but was not one of the targets of the search. The novel condition refers to the case in which the 'target object' did not appear in any of the previous scenes. In order to investigate if some memory for 'target objects' location is formed we compared object's position errors between appeared and novel conditions. Figure 5 displays the mean position errors of all 'target objects' in each condition. It is shown that prior exposure to a 'target object' within a scene reduces the mean position error.

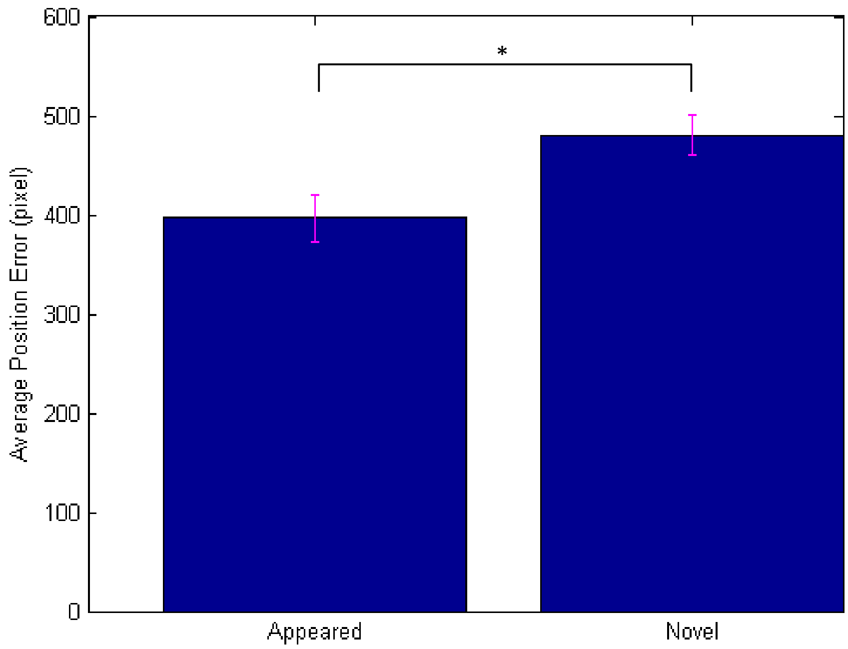

Figure 5. Previous incidental exposure to 'target objects' improves their repositioning. Mean Position error in pixels across all 'target objects' as a function of the two conditions of each object ("appeared" and "novel"). Error bars are standard errors of the participants' position error means.

We also compared position error means of the objects' conditions (appeared vs. novel) between subjects. The results, presented in Table 1, show a significant difference between appeared and novel ($f = 7.128$, $p = 0.014$). Furthermore, we examined the influence of objects' identity and whether they appeared on the position error. Each of these variables individually affects the position error (object identity: f =3.716, p << 0.01; appeared: f = 11.972, p =0.001). The results are seen in Table 2. We can learn from the results that there is a location memory for previously appeared objects. No data points were larger than 4 standard divinations above the mean of each subject.

Table 1. Position error of 'target objects' as a function of object condition (appeared vs. novel)

ANOVA

Dependent Variable: position error

	Sum of Squares	df	Mean Square	F	p
Between Groups	41615.180	1	41615.180	7.128	.014
Within Groups	128438.143	22	5838.097		
Total	170053.323	23			

Table 2. Position error of 'target objects' as a function of object condition (appeared vs. novel) and identity.

ANOVA

Dependent Variable: position error

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Corrected Model	13181427.067 [a]	79	166853.507	2.400	.000	.322
Intercept	92418745.821	1	92418745.821	1329.408	.000	.769
Object	10073921.013	39	258305.667	3.716	.000	.266
Appeared	832303.596	1	832303.596	11.972	.001	.029
Object * Appeared	2275202.458	39	58338.525	.839	.744	.076
Error	27807485.112	400	69518.713			
Total	133407658.000	480				
Corrected Total	40988912.179	479				

a. R Squared = .322 (Adjusted R Squared = .188)

Further investigating the appeared objects, we analyzed how the duration of incidental gaze upon them affected memory of their location.

INCIDENTAL GAZE TIME AND POSITION ERROR

We wanted to test if incidental gaze time on a 'target object' would affect its position error. We plotted the position errors for all 'target objects' of all observers against the sum of all fixation durations on each object that were made during the searches of the 'search objects'. There were no values that were larger than 4 standard deviations above the mean of each subject. Therefore, no data points were excluded. A significant effect was received ($r = -0.1445$, $p < .03$, Figure 6). This negative correlation suggests a memory formation for position of objects as a function of the amount of gaze time spent on an object.

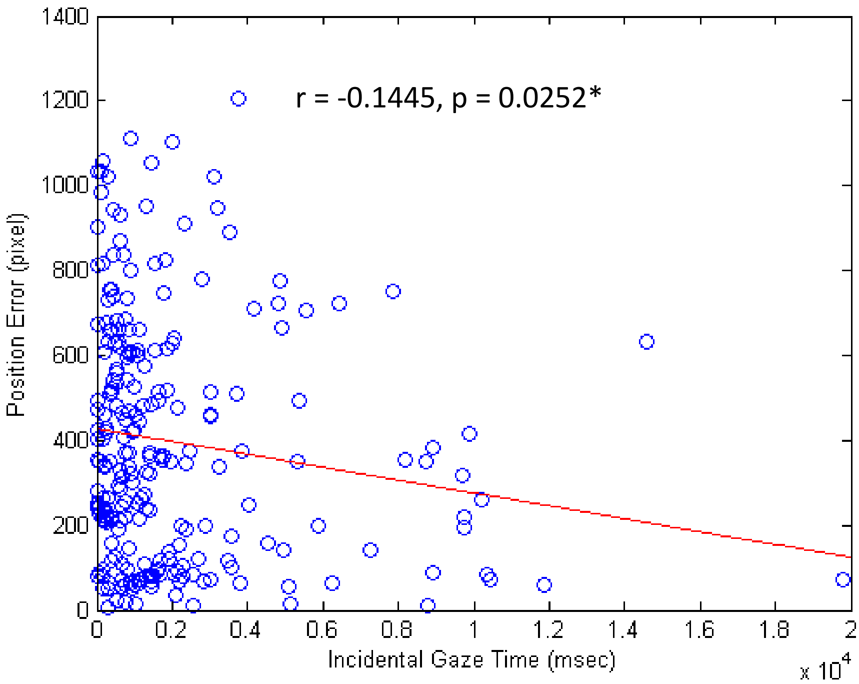

Figure 6. The longer participants were exposed to 'target objects', the more accurate they repositioned them. Position error for all 'target objects' across all participants as a function of the amount of time that a 'target object' was incidentally fixated as distracter on previous search trails. A single data point refers to one object that a subject had located.

OBJECT-LOCATION BINDING IN INCIDENTAL MEMORY:

While a semantic representation of an object was displayed the subjects' answer could lead to 4 different clusters. For the appeared objects: hit and miss; for the novel objects: correct rejection and false alarm (see Table 3). We summarized the answers for each participant and calculated different measures in order to investigate memory performances. Measures were above chance performance (Mean d' = 0.63; see Table 8 in the Appendices). This shows that memory for an object's semantic identity was formed.

Table 3. The 4 possible events' definitions (i.e. clusters) according to the object condition (appeared vs. novel) and the subject's response (yes vs. no).

Subject's Response / Object Condition	"Yes"	"No"
Appeared	Hit	Miss
Novel	False Alarm	Correct Rejection

We wanted to examine if the subjects' answers influenced their location accuracy of the objects. We plotted the mean position errors of all 'target objects' against each cluster. The mean position error received for the hits cluster was the lowest. The results are seen in Figure 7.

There is a significant difference between the mean position error for hits and misses which is shown in Table 4 ($f = 9.626$, $p = 0.002$). No main difference between correct rejections and false alarms was found (see Table 5). In addition, a two-way ANOVA was conducted between subjects on the position error of objects. The independent variables included were object condition (appeared vs. novel) and subjects' answers (correct vs. incorrect). The ANOVA revealed a significant difference between the hits and the other clusters ($f = 9.929$, $p = 0.002$). Namely, the objects that were recognized by the subjects (represented by the hits cluster) were also those for which location memory was most accurate. These results suggest that memory for an object's identity is bound to its location.

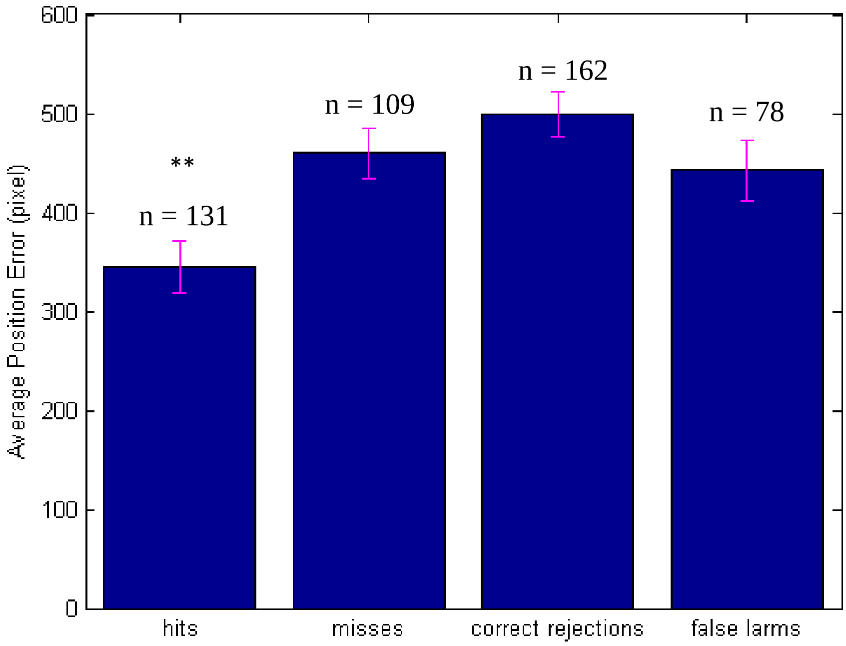

Figure 7. Object-location binding in semantic memory. Mean position error (Euclidean distance in pixels) across all 'target objects' as a function of the 4 clusters. Error bars are standard error of the means. n represents the sample size of each cluster.

Table 4. Position error of 'target objects' as a function of hits vs. misses

ANOVA

Dependent Variable: position error

	Sum of Squares	df	Mean Square	F	p
Between Groups	781108.640	1	781108.640	9.626	.002
Within Groups	19312964.839	238	81146.911		
Total	20094073.479	239			

Table 5. Position error of 'target objects' as a function of correct rejections vs. false alarms

ANOVA

Dependent Variable: position error

	Sum of Squares	df	Mean Square	F	p
Between Groups	168242.981	1	168242.981	2.013	.157
Within Groups	19894292.123	238	83589.463		
Total	20062535.104	239			

SEMANTIC AND VISUAL REPRESENTATIONS

Returning to Figure 3 (1.a-b), the participants were tested on both semantic and visual representations of 'target objects'. In order to test whether the visual hits percentage is above chance level Binomial Test was conducted. The result was positive (p = 0, see Table 6). This indicates that a memory for an object's visual identity is also formed.

Table 6. Binomial Test on visual hits

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (1-tailed)
Visual hits	Group 1	.00	111	.46	.25	.000
	Group 2	1.00	129	.54		
	Total		240	1.00		

In order to examine the influence of semantic encoding and visual encoding on memory of object position, several tests were performed. At first we asked if there are any differences between semantic and visual hits or between semantic and visual misses. Similarity between the mean position errors of semantic and visual representations (can be seen in both hits and misses) is presented in Figure 8. This similarity is probably obtained by a great overlap between hits and misses of the two representations (see Figure 9). Indeed, the probability of a visual hit (A) given that a semantic hit has occurred (B) is very high ($P(A|B) = 0.6565$), and the probability of a visual miss (C) given that a semantic miss (D) has occurred is also very high ($P(C|D) = 0.6055$).

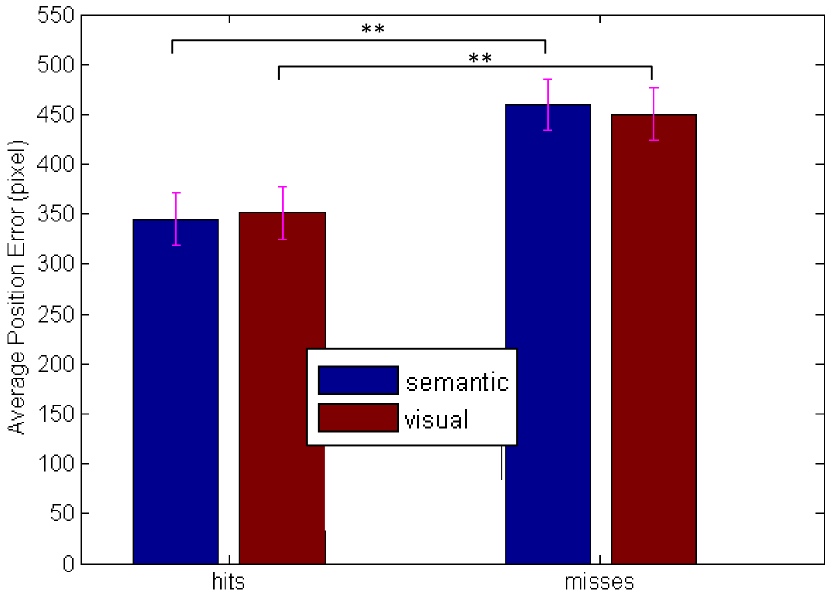

Figure 8. No difference was found between semantic and visual representations for both hits and misses. Mean position error for semantic (blue) and visual (red) representations as a function of hits and misses. Error bars are standard error of the means.

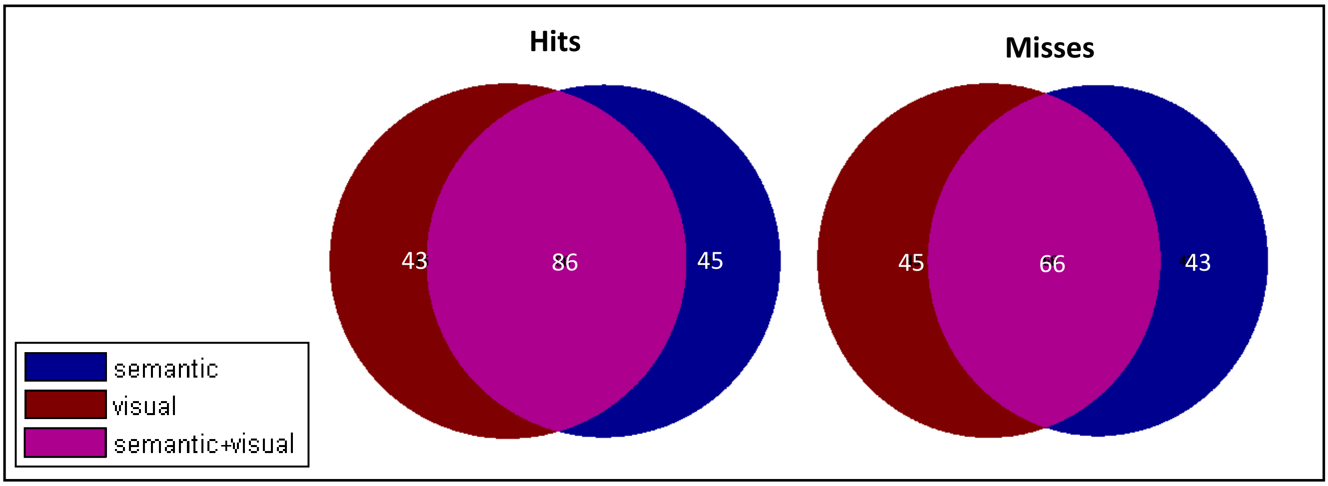

Figure 9. Venn diagrams of the relations between semantic and visual representations. The size of the circles is proportional to the amount of answers in each condition (the numbers are shown in the figure). Red = only visual, Magenta = semantic and visual, Blue = only semantic.

We wanted to examine how encoding (visual and semantic) affects the memory for position of objects. Mean position error was calculated for all the possible events (see Figure 10).

Hit objects both in the semantic test and in the visual test were the most accurately located. The least accurate locations were for miss objects in both tests (left bar and right bar respectively in Figure 10). The middle bars represent the proportional part of each of the encodings. The left middle bar is the mean position error of the objects that were recognized only from their visual representations. The right middle bar represents the mean position error of only the objects that were semantically recognized. By looking at Figure 10 it seems as if semantic encoding and visual encoding have a similar effect on memory for position. The semantic hits and visual hits interaction was not significant (f = 0. 001, p = 0. 972, see Table 7). In addition, no interaction term was found (see Figure 12 in the Appendices).

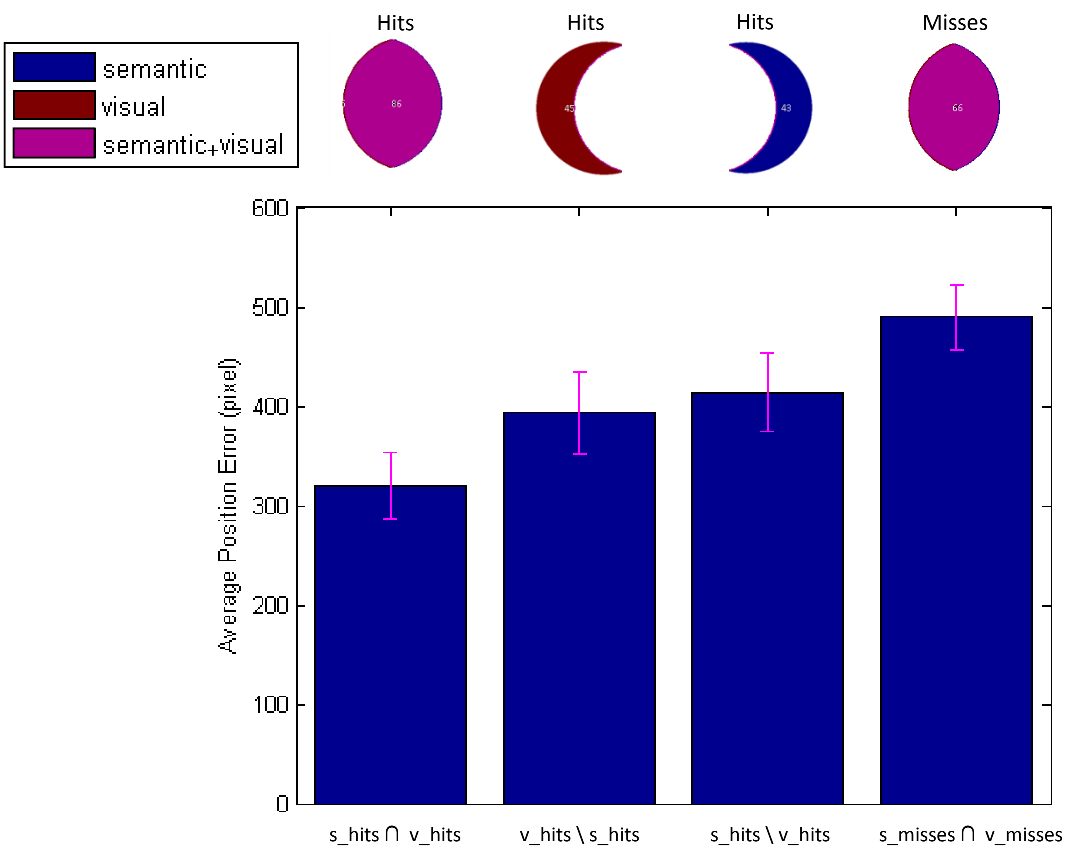

Figure 10. Different types of encoding influence how well position for objects is remembered. Mean position errors as a function of all the possible subsets of hits and misses, between semantic (s) and visual (v). Each bar is corresponding to a specific region on the Venn diagram which are presented above (see also Figure 9). Error bars are standard error of the means.

Table 7. Position error of 'target objects' as a function of subjects' answers (hits vs. misses) and objects' representations (semantic vs. visual).

Tests of Between-Subjects Effects

Dependent Variable: position error

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Corrected Model	1087980.050[a]	3	362660.017	4.503	.004	.054
Intercept	36163575.866	1	36163575.866	449.046	.000	.655
Semantic hits	501307.047	1	501307.047	6.225	.013	.026
Visual hits	306240.810	1	306240.810	3.803	.052	.016
Semantic hits * Visual hits	96.825	1	96.825	.001	.972	.000
Error	19006093.429	236	80534.294			
Total	57949167.500	240				
Corrected Total	20094073.479	239				

a. R Squared = .054 (Adjusted R Squared = .042)

DISCUSSION

After prior search in an unchanging natural scene, participants remembered the semantic and visual properties of objects for which they did not search. They were also able to quite accurately estimate their positions in the scene. Greater incidental gaze time led to improved accuracy of positioning of the 'target objects'. Moreover, position estimates were even more accurate when the 'target object' was encoded both semantically and visually. The ability to code a stimulus in two different ways enables to remember that item better than if the stimulus was only coded in one way (Paivio, 1991). In fact, this shows that long term memory (LTM) for objects and scenes can be quite precise (Hollingworth & Henderson 2002).

It has been previously found that the easier the object is to identify either by shape, color or category, the more likely it will be successfully recognized (Mechelli, et al., 2006, Gerlach, 2009). In our experiment, the visual representation required identification of the detailed structure of the objects, while the semantic required encoding of its general category (e.g. a dog). Therefore, some objects would be easier to encode semantically while other objects would be easier to encode visually. Some categories contain objects which are hard to be distinguished visually from one another, however, they are easier to name. If one sees a clock, for example, one will find it very easy to encode its particular meaning (i.e. category) but may not remember its numeral style (Roman or Arabic) which distinguishes it from other clocks (i.e. its visual representation). Therefore, for this type of objects, it is reasonable to assume that one will find it more difficult to retrieve the object’s visual representation from memory than its semantic representation.

On the other hand, for objects which are very noticeable by shape and color but are hard to name, it is plausible that they would be encoded only visually (Mechelli, et al., 2006, Gerlach, 2009).

In our experiment, we were able to distinguish between the influences of semantic encoding and visual encoding on how well position for objects is remembered. A resembling effect was received for both types of encoding (Figure 10).

Mishkin & Ungerleider (1982) hold that the neural areas supporting object recognition are separate from those supporting location encoding (Humphreys, 1998). If encoding the position of an object is separate from recalling it, the similar effect of both visual and semantic encoding on the position error can be explained. As was mentioned before, the ability to code a stimulus in two different ways enables to remember that item better. Encoding position of an object combined with at least one of the codes (visual/ semantic) can achieve better performance in retrieving its position from memory.

Similar to previous findings by Võ & Wolfe (2012, 2011) and by Oliva et al. (2004), we found that repeatedly searching for multiple different objects in the same unchanging scene does not dramatically speed up the search. However, it was previously found that memory for object is formed during search for other objects (Võ & Wolfe, 2012). In our experiment participants were able to locate and recognize the ‘target objects’ which indicates that memory, for these objects, was formed. This raises a question: why did not the participants use their memory during the search task such that the search for the latter objects in the same scene was faster than the first objects due to scene familiarity? A possible explanation might be that while the scene is displayed it may be faster to simply search again online rather than to retrieve and reactivate stored memory traces. In contrast, in the memory test the subjects were forced to retrieve and reactivate stored memory traces of the 'target objects'. Our results demonstrate incidental memory for objects and their positions and indicate that memory for an object's identity is bound to its location.

One of the main limitations in the experiment is that the scenes are presented in 2D while they are simulating a 3D environment. Thus, the method in which the position error for object is calculated can yield misrepresentative results. For example, an object that is located in midair can be accounted as a small position error. Moreover, one can reposition the aquarium In Figure 11 on the floor (B) and it can be accounted as more accurate than another repositions on the table (C), although the second position is more accurate because the aquarium indeed was on the table (A). In order to cope with this difficulty, we will adjust the defined areas of location (i.e. table).

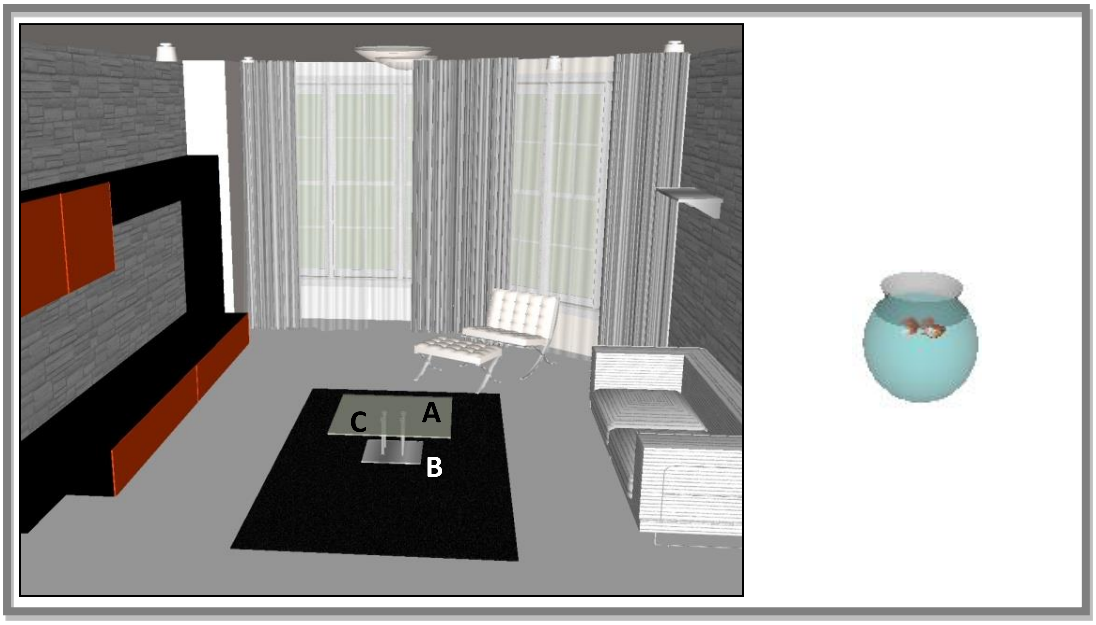

Figure 11. The aquarium as the 'target object' in the position task is on the right and on the left is its original scene. The position possibilities of the aquarium, shown in the figure, are: the correct position of the aquarium (A), positioning the aquarium on the floor (B), positioning the aquarium on the table but not in the correct location (C).

Taken together, our findings demonstrate that binding to location in visual and semantic memory is acquired intentionally. Indeed, this study serves as preliminary results toward subsequent research. The next challenge is to examine possible reasons for the similar effects (position error magnitude) of the visual and semantic encodings.

ACKNOWLEDGEMENTS

Great thanks to Nomi Carlebach for her extraordinary support in this project process, Dr. Tanya Orlov for consulting the statistical analysis, Yuval Porat, Dr. Ayelet McKyton and Zvi Roth for their assistance. Finally I would like to express my special appreciation and thanks to my advisor Professor Ehud Zohary.

REFERENCES

Boot, W. R., McCarley, J. M., Kramer, A. F., & Peterson, M. S. (2004). Automatic and intentional memory processes in visual search. *Psychonomic Bulletin & Review, 5*, 854–861.

Castelhano, M. S., & Henderson, J. M., (2005). Incidental visual memory for objects in scenes. *Visual Cognition, 12*, 1017–1040.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*(5), 170-178.

EyeLink 1000 Technical specifications, 2013 SR Research (http://www.sr-research.com/pdf/techspec.pdf).

Gerlach, C. (2009). Category-specificity in visual object recognition. Cognition, 111, 281-301.

Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology, 10*, 1209–1212.

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 210–228.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology-Human Perception and Performance, 281*, 113-136.

Hollingworth, A. (2005). Memory for object position in natural scenes. *Visual Cognition*, *12*, 1003-1016.

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*, *17*, 273–291.

Hollingworth, A. (2012). Guidance of visual search by memory and knowledge. In M. D. Dodd & J. H. Flowers (Eds.), *The Influence of Attention, Learning, and Motivation on Visual Search, Nebraska Symposium on Motivation* (pp. 63-89). New York: Springer.

Humphreys, G. W. (1998). Neural representation of objects in space: a dual coding account. *Philos Trans R Soc Lond B Biol Sci., 353*(1373): 1341–1351.

Kellenbach, M., Hovius, M., & Patterson, K. (2005). A PET study of visual and semantic knowledge about objects. *Cortex, 41*, 121-132

Klein, R. M. (2009). On the control of attention. Canadian *Journal of Experimental Psychology*, *63*, 240–252.

Kristjansson, A. (2000). In search of remembrance: Evidence for memory in visual search. *Psychological Science*, *11*, 328–332.

Körner, C., & Gilchrist, I. D. (2007). Finding a new target in an old display: Evidence for a memory recency effect in visual search. *Psychonomic Bulletin & Review, 14*, 846–851.

Körner, C., & Gilchrist, I. D. (2008). Memory processes in multiple-target visual search. *Psychological Research, 72*, 99–105.

Mechelli, A., Sartori, G., Orlandi, P., & Price, C. (2006). Semantic relevance explains category effects in medial fusiform gyri. *NeuroImage*, *30*, 992-1002.

Mishkin, M., Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav Brain Res, 6*(1): 57–77.

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research, 46*(5), 614–621.

Oliva, A., Wolfe, J. M., & Arsenio, H. C. (2004). Panoramic search: The interaction of memory and vision in search through a familiar scene. *Journal of Experimental Psychology: Human Perception and Performance, 30*(6), 1132–1146.

Paivio, Allan (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *45*(3), 255-287.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science, 12*, 287–292.

Peterson, M. S., Beck, M. R., & Vomela, M. (2007). Visual search is guided by prospective and retrospective memory. *Perception & Psychophysics, 69*, 123–135.

Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J. V., Fadiga, L., & Pietrini, P., (2009). Do we really need vision? How blind people "see" the actions of others. *J Neurosci, 5*; 29(31):9719-24.

Sternberg, R. J., & Grigorenko, E. L., (2003). The Psychology of Abilities, Competencies, and Expertise. *Cambridge: Cambridge University Press*.

Takeda, Y., & Yagi, A. (2000). Inhibitory tagging in visual search can be found if search stimuli remain visible. *Perception & Psychophysics, 62*,927–934.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766–786.

Võ, M. L. H., & Wolfe, J. M. (2012). When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. J. Exp. Psychol: Human Perception and *Performance, 38*(1), 23-41.

Võ, M. L. H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition, 126*(2), 198-212.

Võ, M. L. H., Schneider, W. X., & Matthias, E. (2008). Transsaccadic scene memory revisited: A 'Theory of Visual Attention (TVA)' based approach to recognition memory and confidence for objects in naturalistic scenes. *Journal of Eye-Movement Research, 2*(2), 1–13 (Article no. 7).

Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and non-selective pathways. *Trends in Cognitive Sciences, 15*(2), 77–84.

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research, 49*(16), 2095–2103.

APPENDICES

Table 8. For each subjects the following variables that indicate on memory formation were calculated: d', beta, semantic variables (hit, miss, correct rejection, false alarm) and visual variables (hit, miss).

			Semantic (%)				**Visual (%)**	
Subject	**d'**	**beta**	**hit**	**miss**	**correct rejection**	**false alarm**	**hit**	**miss**
1	0.25	0.97	60	40	50	50	55	45
2	0.78	1.66	40	60	85	15	55	45
3	1.20	0.91	75	25	70	30	55	45
4	0.64	0.96	65	35	60	40	45	55
5	0.40	1.14	45	55	70	30	60	40
6	1.39	3.75	40	60	95	5	55	45
7	0.78	1.11	60	40	70	30	35	65
8	0.39	0.93	65	35	50	50	40	60
9	1.05	1.00	70	30	70	30	70	30
10	0.13	0.99	55	45	50	50	55	45
11	0.38	0.98	60	40	55	45	65	35
12	0.19	1.20	20	80	85	15	55	45
Average	0.63	1.30	54.58	45.42	67.50	32.50	53.75	46.25

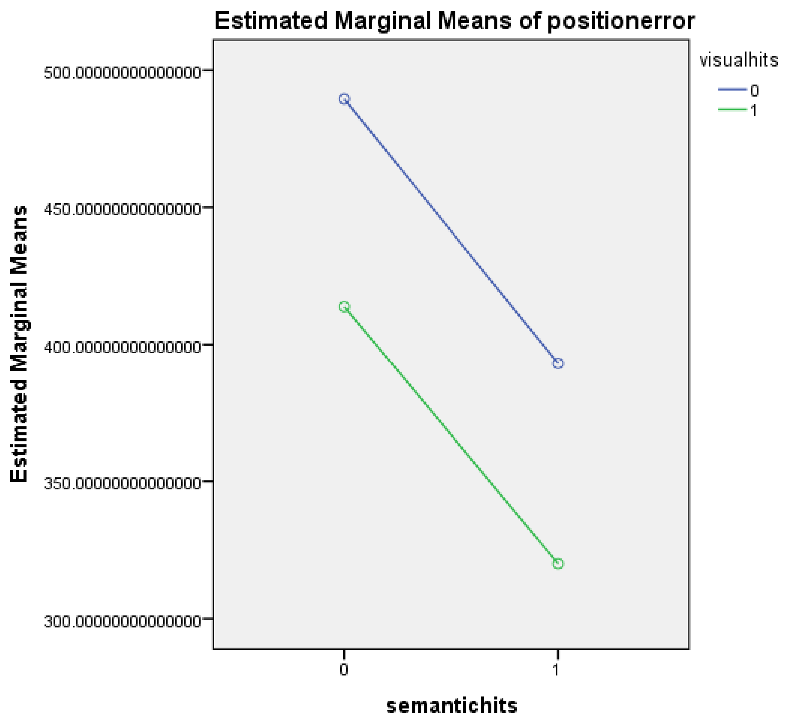

Figure 12. Multiple Regression – Interaction test between visual hits and semantic hits.