

# Of “what” and “where” in a natural search task: Active object handling supports object location memory beyond the object’s identity

Dejan Draschkow<sup>1</sup> · Melissa L.-H. Võ<sup>1</sup>

Published online: 10 May 2016  
© The Psychonomic Society, Inc. 2016

**Abstract** Looking for as well as actively manipulating objects that are relevant to ongoing behavioral goals are intricate parts of natural behavior. It is, however, not clear to what degree these two forms of interaction with our visual environment differ with regard to their memory representations. In a real-world paradigm, we investigated if physically engaging with objects as part of a search task influences identity and position memory differently for task-relevant versus irrelevant objects. Participants equipped with a mobile eye tracker either searched for cued objects without object interaction (Find condition) or actively collected the objects they found (Handle condition). In the following free-recall task, identity memory was assessed, demonstrating superior memory for relevant compared to irrelevant objects, but no difference between the Handle and Find conditions. Subsequently, location memory was inferred via times to first fixation in a final object search task. Active object manipulation and task-relevance interacted in that location memory for relevant objects was superior to irrelevant ones only in the Handle condition. Including previous object recall performance as a covariate in the linear mixed-model analysis of times to first fixation allowed us to explore the interaction between remembered/forgotten object identities and the execution of location memory. Identity memory performance predicted location memory

in the Find but not the Handle condition, suggesting that active object handling leads to strong spatial representations independent of object identity memory. We argue that object handling facilitates the prioritization of relevant location information, but this might come at the cost of deprioritizing irrelevant information.

**Keywords** Real-world search · Mobile eye tracking · Action · Object handling · Eye movements · Object memory · Incidental memory · Task influences

## Introduction

In natural behavior, cognitive processes are strongly intertwined with observers’ active interaction with the environment. Already in 1979, Gibson proposed that affordances connected to completing an action certainly play an important role in the processing of our immediate environment (Gibson, 1979). The notion that cognition should be investigated as a function of available actions has since become increasingly popular (Clark, 1999; Engel, Maye, Kurthen, & König, 2013; McGann, 2007). One particular branch of research has focused on an action-specific account for perception, postulating that our perception of the environment is contingent on the perceiver’s capabilities to act and interact with the environment (e.g., Witt & Riley, 2014; Witt, 2011). In our study, using real-world furnished rooms, we investigated how acting on objects influences location and identity memory for these objects and what role task relevance plays in the formation of object representations.

Object identities are recognized faster after active rotation compared to passive rotation observation (Harman, Humphrey, & Goodale, 1999; James et al., 2002). Active hand movements to locations of abstract objects lead to better recall

---

✉ Dejan Draschkow  
draschkow@psych.uni-frankfurt.de; <http://www.draschkow.com>;  
<http://www.SceneGrammarLab.com>

<sup>1</sup> Scene Grammar Lab, Department of Cognitive Psychology, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 6, Frankfurt am Main 60323, Germany

of the locations and faster reaction times to the locations than when the hand is passively moved (Trewartha, Case, & Flanagan, 2015). Picking up and holding isolated objects modulates the spatial representation of the targets (Thomas, Davoli, & Brockmole, 2012). A study comparing active engagement in a natural task (preparing tea) with first-person observation of the same manipulations via video (gaze recordings of tea preparation) found that physical manipulation of relevant objects (tea-related) results in prioritization for object position memory (Tatler et al., 2013). Spatial memory for task-irrelevant objects was deprioritized and did not differ from chance in the natural task, but was above chance in the first-person observation condition, hinting towards an interactive relationship between the relevance of objects to one's own ongoing behavior and the physical manipulation of these objects.

Contrary to these findings, there are data indicating that in some respects active behavior does not enhance memory representations. Even though recognition time for objects was enhanced in the active condition of their study, there was no difference in accuracy between active and passive manipulation (Harman et al., 1999; James et al., 2002). Recall and recognition memory after active exploration of a virtual environment was improved for neither object identity nor location memory – instead only the spatial layout of the virtual reality environment improved (Brooks, Attree, Rose, Clifford, & Leadbetter, 1999). No prioritization of location memory was found for physically manipulated objects in a natural task, nor for relevant compared to irrelevant ones (Kirtley & Tatler, 2015). In sum, these contradictory findings leave the question of the impact of active object handling on identity and location memory, as well as its interaction with the relevance of objects to ongoing behavior unresolved. In order to settle this open question, we sequentially tested both identity and location memory after either active or passive interaction while additionally maintaining a task-relevant and irrelevant distinction of objects.

The focus on the studies presented so far was on the distinction between passive observation of an action and more natural proactive engagement with our environment. Human vision, however, is a dynamic and active process and is needed to support ongoing behavior goals (Findlay & Gilchrist, 2001; Henderson, 2003, 2007). Active search not only influences our gaze behavior (Castelhano, Mack, & Henderson, 2009), but actively looking *for* an object embedded in a scene actually boosts location memory for the same object compared to simply looking *at* the object (Vö & Wolfe, 2012). Incidental encoding during search in scenes even leaves participants with better object identity memories than intentional memorization of these (Draschko, Wolfe, & Vö, 2014). The availability of a structured and meaningful scene context might therefore not only support object search (for a review see Vö & Wolfe, 2015), but also the formation of memory

representations highlighting the importance of research in natural environments. In our current study, participants had to complete an active search task within a real-world environment while we manipulated if the participants interacted with an object or not.

In real-world interactions with the environment, strictly observational behavior is rare. The goal of our study was therefore to investigate the influence of object handling on object memory within overall active tasks. Using a fully furnished four-room apartment as a real-world environment, participants were either asked to search for (Find condition) or collect (Handle condition) objects. So in comparison to previous work, both experimental conditions reflected natural, active behavior. A surprise free-recall task followed, in which participants were asked to recall all remembered objects. This measure of identity memory was later used in our linear mixed modelling analysis as an additional predictor for the subsequent location memory test, as identity and location memory are strongly related (Olson & Marshuetz, 2005). At the end of the experiment, participants' location memory was inferred by a repeated search for the initial objects, providing a natural way of testing extant spatial representations. During the whole study, participants were equipped with a mobile eye tracker, as memory performance can be predicted by gaze durations (Hollingworth & Henderson, 2002) and number of fixations (Tatler, Gilchrist, & Land, 2005; Tatler & Tatler, 2013). Moreover, this provided us with a more fine-grained measure of search times, i.e. the time to first fixation of the target object.

To anticipate our main findings, we found that location memory was modulated by an interaction of task relevance and object handling, whereas identity memory was not influenced by object handling. Relevant objects which were manipulated were found faster than objects which were not handled, but this was not true for irrelevant ones. Identity memory performance predicted location memory for non-handled objects, while location memory for actively handled objects was independent of recall performance. In combination, these findings suggest that active object handling leads to prioritization of relevant and deprioritization of irrelevant location information, above and beyond identity memory performance.

## Methods

### Participants

Sixteen participants (mean age = 20.7 years, range = 18–26, 11 female, 15 right-handed) were recruited at the Goethe University Frankfurt. All had normal or corrected-to-normal vision. All were volunteers receiving course credit and had given informed consent.

## Apparatus

Participants' eye movements were tracked at a sampling rate of 60 Hz using the SMI Eye Tracking Glasses, which allow for the recording of both eyes with automatic parallax compensation with a spatial accuracy of about 0.5°. The scene video was recorded at a resolution of 960 × 720 at 30 frames/s, with a field of view of 60° (horizontal) and 46° (vertical). Event detection was performed offline with BeGaze software by SMI. Calibration involved asking participants to look at three different objects allocated at different distances, reflecting the depth of the experimental rooms. Sound was recorded throughout the experiment. Auditory cues were presented with MATLAB, 2012b using Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on a laptop running Windows 7.

## Environment

Four rooms (each approximately 15 m<sup>2</sup>) of a real-world apartment (approximately 67 m<sup>2</sup>) in Frankfurt, Germany constituted the experimental environment of this study (Fig. 1, top). The fifth room, a bathroom, was used for practice trials. All rooms were carefully prepared to represent their semantic category (kitchen, bedroom, study, living room) containing objects that are typical for the specific type of room. Further, object locations were chosen to be as typical as possible to avoid potential effects of unusual placement on inspection behavior (Vö & Henderson, 2009). The environment was the same for all conditions and participants.

Each room contained an a priori defined set of 15 task-relevant and 15 task-irrelevant objects (overall 120 objects; see Fig. 1, bottom). The relevance of each object was determined by the participant during the experiment in accordance with the instruction described in the “**Procedure**” section below. Overall this led to 51 % of all objects being categorized as relevant and 49 % as irrelevant. Eight percent of all trials were categorized by the participants differently than our a priori selected relevance. The rest of the interior consisted of furniture and background objects congruent with the room's category.

## Procedure

Upon arrival, participants were informed that they would be searching for objects in the apartment and that in some cases they would have to simply respond verbally that they had found an object (Find condition), while in other cases they would have to additionally collect the target (Handle condition). The written instructions also included a description of relevant object categories: Participants were told to imagine going on a ski trip and had to indicate for each object if they would want to pack the given object or not. This was done either only verbally (Find condition) or by additionally

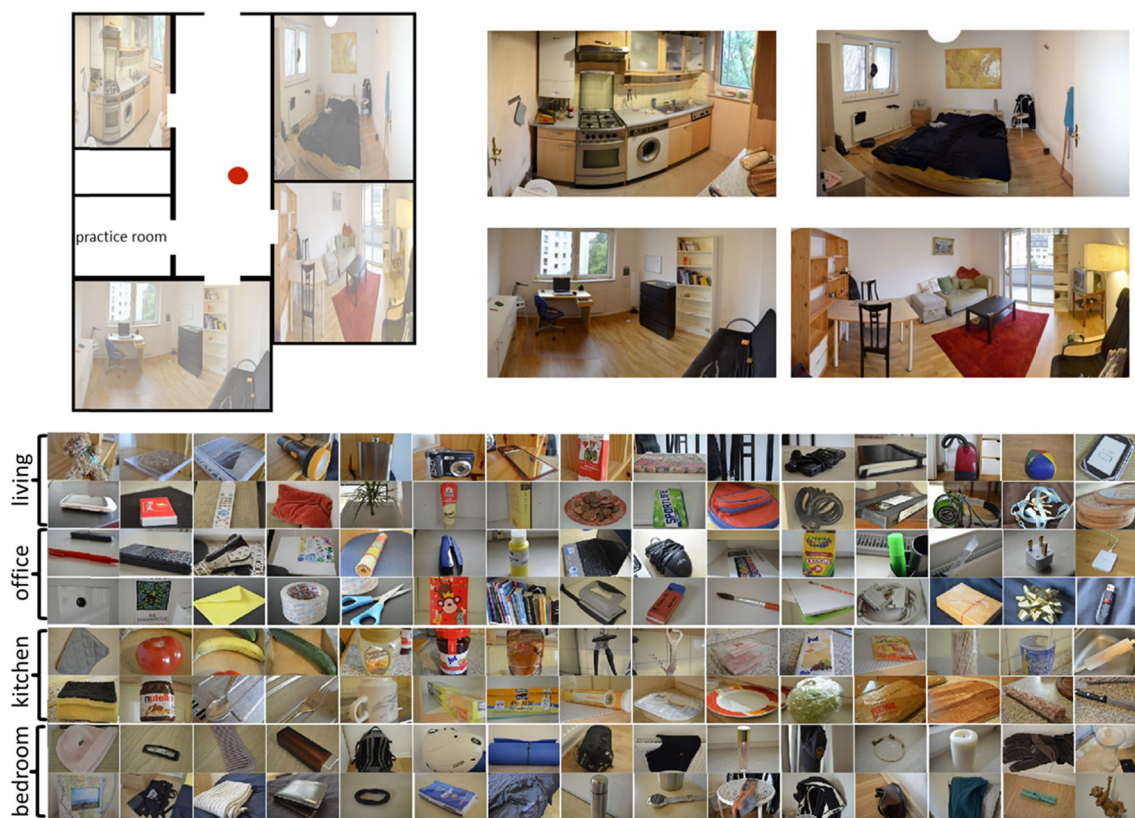
placing the object into a container (Handle condition). The instructions further informed them that only objects from a certain category in each room would be important for the trip: kitchen – objects needed for preparing a sandwich, bedroom – ski equipment and clothing, study – objects needed for a guest present, living room – objects related to entertainment. The instructions were repeated verbally to the participant after the mobile eye tracker was calibrated and six practice trials were completed.

The experimental session lasted about 75 min and consisted of three phases: the Packing phase, the Free-recall phase, and the Search phase.

**Packing phase** Participants were encouraged to move around and explore the room freely during search in order to find the object as quickly as possible. Each trial began with the participant standing in the hallway of the apartment (as indicated with the red dot in Fig. 1) in front of a fixation cross, with no visibility of any room interior. An auditory cue informed the participant of the current task (Handle: “Collect...” vs. Find: “Find...”), the target object (“toothbrush”), as well as the room the object was to be found in (“bathroom”, see Fig. 2, top). In both tasks the participant responded verbally (“Yes!”) as soon as they found the object. In Find trials, the participant would return to the hallway after responding. In Handle trials, the participant would collect the target object before returning and thus physically interact with the object. Subsequently, in both conditions, the participant would indicate verbally if the object of the current trial should be packed or not (Relevant vs. Irrelevant). Additionally, when the search was active, the participant had to put the object in a covered suitcase (Relevant) or in a covered box (Irrelevant), which were placed on the floor below the fixation cross. The order of searches was randomized for all participants across objects and rooms. The assignment of objects to the Handle (40 trials) and Find (40 trials) conditions was counterbalanced across participants. Distractor objects in the rooms remained the same for all participants. This phase consisted of 80 trials. Video material of example trials is available on <http://www.scenegrammarlab.com/research/real-world/of-what-and-where-in-a-natural-search-task/>

**Free-recall phase** After completing the Packing phase, participants were taken out of the environment for a surprise recall task in which they were asked to write down the name of every object they remembered from the four experimental rooms on a sheet of paper (Fig. 2, middle). They were specifically instructed to not only recall objects they had searched for, but also objects in the environment with which they had never interacted.

**Search phase** While participants were busy writing down object names from memory, the experimenters rearranged



**Fig. 1** Depiction of the experimental environment used in this study. The top left figure shows a simplified ground plan of the apartment and the starting position of the participant in each trial is indicated with a red dot.

On the top right panoramic photographs of the four experimental rooms are depicted. The bottom images show the 120 critical objects used in the study

the environment to its original state at the start of the experiment. The participants received new instructions for this last phase and were then asked back into the apartment after completing the recall task. According to the instructions, objects from the first Packing phase would have to be found again in a trial-by-trial manner, as well as other objects that were in the rooms, but had not yet been searched for (Distractors). Additionally, participants were informed that some objects would not be present in any of the rooms (Absent). As in the Packing phase, each trial began with the participant standing in the hallway of the apartment looking at the fixation cross (Fig. 2, bottom).

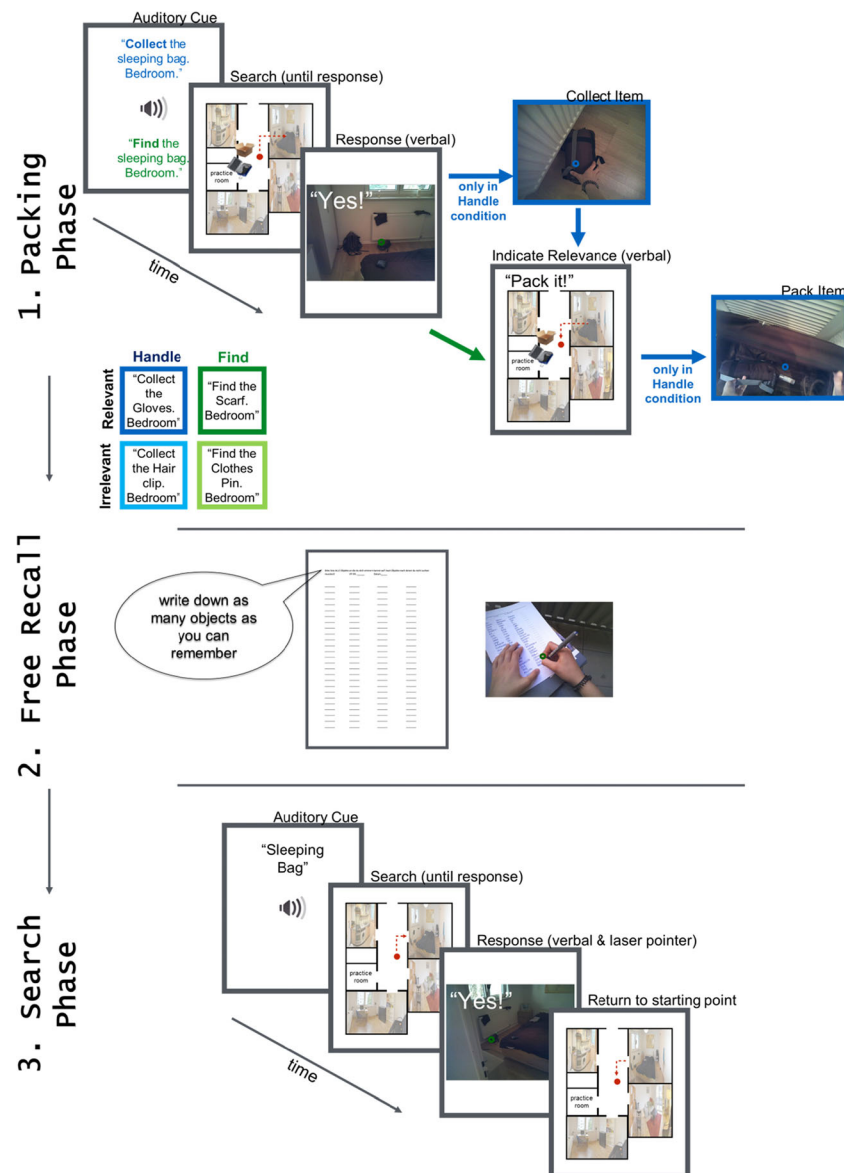
An auditory word cue indicated the target object of the current trial. The participants responded verbally (“Yes”/”No”) as soon as the object was found or determined not present. Contrary to the Packing phase, participants had to illuminate the target object with a laser pointer as soon as the object was found. They were instructed to complete the searches as fast and as accurately as possible and then return to the fixation cross after responding. The order of searches for previously Handle (40 trials) and Find (40 trials), as well as Distractor (40 trials) and Absent (40 trials) objects was randomized for all participants across objects and rooms. This final phase consisted of 160 trials.

**Data analysis**

For eye movement analyses, event detection was performed offline and videos were created showing gaze estimation for each fixation with the software BeGaze by SMI. Eye-tracking data were coded using the *Semantic mapping* option in BeGaze. This software automatically detects candidate fixations and allows each to be tagged manually with the corresponding interest area.

A search was deemed successful when the participant responded to the right target object, as validated from the gaze recordings. Total gaze durations for each object were calculated by summing up the time spent fixating an object throughout a phase. Total *incidental* gaze durations only included the fixation duration on objects before they became a target in each experimental phase. For both the Packing and Search phases reaction time was defined as the time from the end of the auditory cue until the participant’s verbal response. Time to first fixation was the time from the end of the auditory cue until the first fixation on the target object of the current trial.

For analyzing the effects in our data, linear mixed-effects models (LMM) were run using the lme4 package (Bates,



**Fig. 2** Examples of the trial procedure for the Packing phase (top) and the Search phase (bottom). Example recall sheet for the Free-recall phase is depicted in the middle

Maechler, Bolker, & Walker, 2014) in the R statistical programming environment (R Development Core Team, 2012). Contrasts were defined to analyze the critical comparisons. For continuous dependent variables, confidence intervals (CIs) for model parameters were obtained via profiling, using the *confint* function in R. For binary responses, p-values based on asymptotic Wald tests were computed within the framework of generalized linear mixed-effects models. We chose the LMM approach as it allows between-subject and between-item variance to be estimated simultaneously and thus yields potential advantages over traditional F1\F2 analysis of variance (for a discussion see Baayen, Davidson, & Bates, 2008; Kliegl, Wei, Dambacher, Yan, & Zhou, 2010). Participants and rooms were included in the models as random factors. In practice, models with random intercepts and slopes

for all fixed effects often fail to converge or lead to over-parameterization (Bates, Kliegl, Vasishth, & Baayen, 2015). In order to produce models that converge on a stable solution and are properly supported by the data, we used a drop-one procedure starting with the full model including all varying intercepts and varying slopes of the main effect of the experimental design. Varying slopes not contributing significantly to the goodness of fit (likelihood ratio tests) were removed from the model. Following inspection of the distribution and residuals, gaze duration and time to first fixation were log-transformed in order to more closely approximate a normal distribution and meet LMM assumptions. Details about the individual analysis and models are described in the “Results” section.

## Results

### Packing phase – influences of incidental fixations

Reaction times and total incidental gaze durations exceeding the individual's mean by  $\pm 2$  standard deviations were excluded from analysis. The reaction time criterion led to the removal of 4.3 % of the trials and the outlying gaze durations led to an additional removal of 3.6 % of the data. Error trials were removed as well (5.4 % misses), but incidental fixations during these trials were included in the analysis, as exposure to objects during error trials still contributes to the formation of object representations. Error trials mainly consisted of trials in which either an object was mistakenly collected (Handle) in a Find trial or an object was not collected (Find) in a Handle trial.

**Repeated search in the same environment** For the search time model, the retained variance components in the best fitting and converging LMM were the item (rooms) and subjects intercepts. Time to first fixation did not differ between the Handle and the Find condition,  $t < 1$  and there was no difference between relevant and irrelevant targets,  $t = -1.03$  (Fig. 3, left). The interaction between the two was also not significant,  $\beta = 0.054$ ,  $SE = 0.028$ ,  $t = 1.88$ , 95 % CI  $[-0.003, 0.108]$ . Participants improved in their search time in general over trials,  $\beta = -0.002$ ,  $SE = 0.0001$ ,  $t = -5.69$ , 95 % CI  $[-0.002, -0.001]$ , but this improvement did not differ between Handle and Find trials,  $t < 1$ , or between relevant and irrelevant ones,  $t < 1$ .

**The role of incidental fixations** We used total incidental gaze durations on an object prior to becoming a target to test whether fixating a distractor object during a search for another target would improve search time once this object becomes a target. The summed incidental gaze durations on each object before it became the target of the current trial were correlated with the time to first fixation for that object (Fig. 3, right). A regression line fitted to the data had a strong negative slope, but was not significantly different from zero ( $\beta = -0.187$ ,  $SE = 0.160$ ,  $t = -1.17$ ,  $p = 0.24$ ). The disadvantage of this approach is that it does not consider between-subject and between-item differences which might be of significant importance not only in real-world but also in computer-based paradigms (Kliegl et al., 2010). To counter the disadvantage of merely fitting a regression line, we included summed fixation durations as a covariate in the initial model for time to first fixation. Model outputs indicate that total incidental gaze durations significantly predicted time to first fixation,  $\beta = -0.063$ ,  $SE = 0.026$ ,  $t = -2.46$ , 95 % CI  $[-0.071, -0.017]$ , with higher gaze durations leading to faster search times. Crucially, the inclusion of total incidental gaze durations as a covariate eliminated

the previously significant effect of trials on search time,  $\beta = -0.0004$ ,  $SE = 0.0004$ ,  $t = -1.05$ , 95 % CI  $[-0.070, 0.0003]$ .

This result dissociates the role of multiple exposures to the same environment and total incidental gaze durations on objects embedded in that environment. Repeatedly searching through the same environment by itself does not yield any benefits for search time, and contextual information of the environment might be sufficient for guiding search behavior. Only in combination with an accumulation of incidental gaze durations on future search targets can repeated exposure lead to a performance improvement.

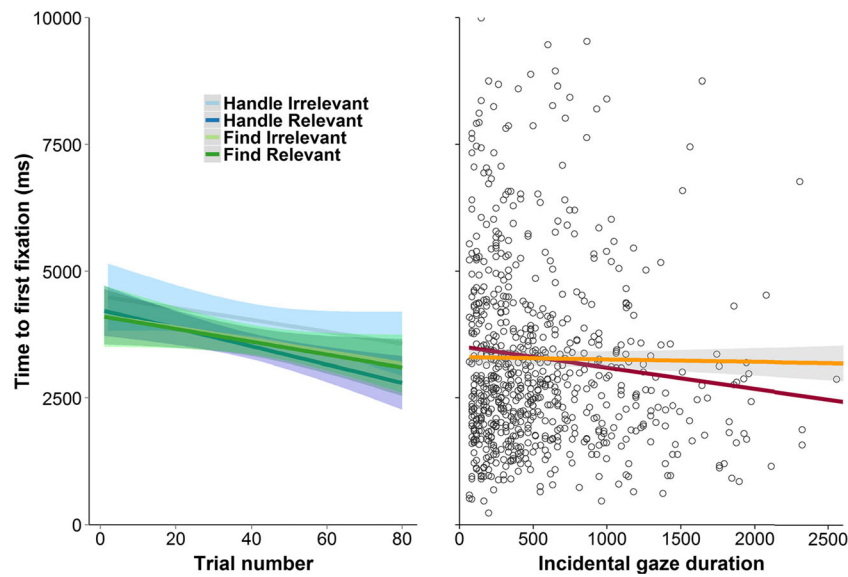
### Free-recall phase – identity memory assessment

For the identity memory model, the set of random components retained in the best fitting LMM were the item (rooms) and subject intercepts and the effect of the experimental condition (Handle vs. Find vs. Distractor) for subjects. Targets (objects searched for in the Handle and Find conditions) were remembered better than Distractors (objects which were embedded in the rooms, but never searched for)  $\beta = -2.291$ ,  $SE = 0.238$ ,  $z = -9.645$ ,  $p < 0.01$  (Fig. 4, left). Recall performance for task-relevant objects exceeds that for irrelevant ones in all conditions,  $\beta = 0.513$ ,  $SE = 0.130$ ,  $z = 3.934$ ,  $p < 0.01$ . There was no difference between the Handle and the Find condition,  $\beta = -0.122$ ,  $SE = 0.131$ ,  $z = -0.934$ ,  $p > 0.3$ .

**Including fixation durations from the Packing phase as a covariate** Gaze durations from the Packing phase were summed for each object and were included in the model as a covariate. Gaze durations on objects during the Packing phase did not predict identity memory,  $\beta = 0.215$ ,  $SE = 0.206$ ,  $z = 1.044$ ,  $p > 0.2$ . The inclusion of the covariate did not change the significance of the critical comparisons presented above.

### Search phase – assessing location memory

Reaction times and total incidental gaze durations in the Search phase exceeding the individual's mean by  $\pm 2$  standard deviations were excluded from analysis. This reaction time criterion led to the removal of 4.7 % of the data and the criterion for gaze duration to the exclusion of an additional 1.8 %. Trials in which participants responded that an object was absent when in fact it was present (mean misses for objects which were previously in the Handle = 6.1 %, Find = 10.1 %, and Distractor = 33.7 % condition) or responded that an object was present when in fact it was absent (false alarms = 6.2 %), were considered as error trials. Trials in which the participants had to search for objects which were erroneously searched for/handled during the preceding Packing phase were removed from the data. Together with error trials this led to the removal of 13.9 % of trials.



**Fig. 3** Left: Time to first fixation as a function of the trial number and the condition. Continuous error bars (shown in the appropriate color for each condition) are 95 % within-subject confidence intervals. Right: Time to first fixation for each object and participant as a function of the amount of

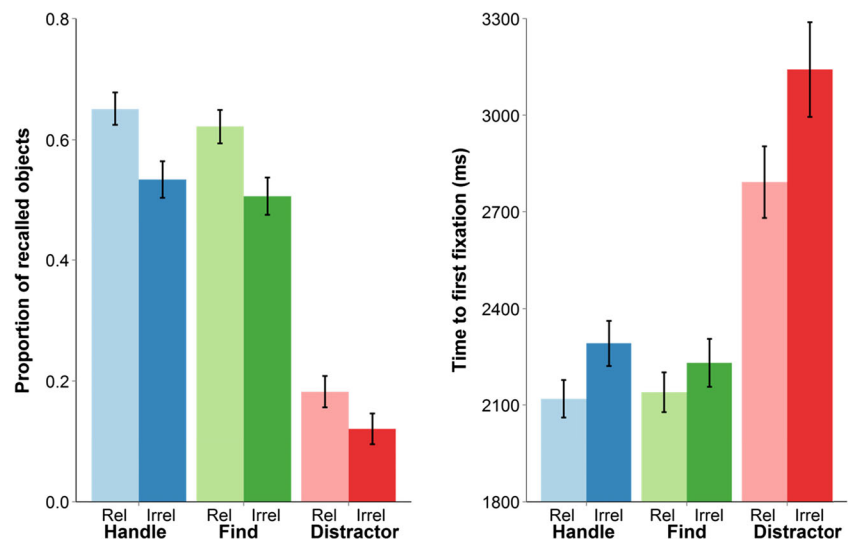
time the target was incidentally fixated on previous trials. The orange regression line is fitted to all data points, disregarding participant and item variation. The red line represents the slope of the appropriate linear mixed-effects model

There was a general effect of target-present trials being faster than target-absent trials,  $F(1, 15) = 48.71, p < 0.01$ . Search time analysis was conducted only with target-present trials.

CI  $[-0.040, -0.005]$ . There also was a significant difference between relevant and irrelevant Distractors,  $\beta = -0.02, SE = 0.011, t = -2.16, 95\% \text{ CI } [-0.044, -0.002]$ .

The retained variance components for the final best fitting location memory model were the item and subject intercepts. Time to first fixation was faster for Targets than Distractors,  $\beta = 0.120, SE = 0.012, t = 9.68, 95\% \text{ CI } [0.096, 0.145]$  (Fig. 4, right). There was no search time difference between the Handle and the Find condition,  $t < 1$ . There was no effect of Relevance on time to first fixation in the Find condition,  $t < 1$ , but relevant objects were fixated faster than irrelevant ones in the Handle condition,  $\beta = -0.02, SE = 0.009, t = -2.58, 95\%$

**Including ALL fixation durations during the experiment as a covariate** Gaze durations from the Packing phase as well as total incidental gaze durations on objects in the Search phase prior to them becoming a target were summed for each object and were included in the model as a covariate. Summed fixation durations on objects during the experiment predicted search time,  $\beta = -0.115, SE = 0.020, t = -5.77, 95\% \text{ CI } [-0.116, -0.042]$  in that longer gaze durations led to shorter search times. With the



**Fig. 4** Memory performance as a function of the experimental conditions for relevant (Rel) and irrelevant (Irrel) objects. Left: The proportion of recalled objects during the Free-recall phase. Right: Time to first fixation during the Search phase

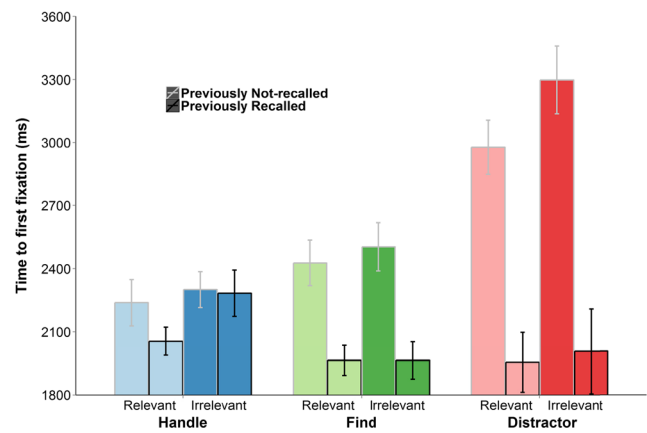
inclusion of fixation durations as covariate, the search time difference between relevant and irrelevant Distractors was not significant,  $\beta = -0.018$ ,  $SE = 0.012$ ,  $t = -1.55$ , 95 % CI  $[-0.041, 0.005]$ , indicating that the initial effect was driven by the difference in fixation durations for relevant versus irrelevant Distractors. There was no change in significance for the rest of the critical comparisons.

### Dissociating between “what” and “where”

In order to investigate if location memory — measured as times to first fixation in the final Search phase — was predicted by identity memory, we included recall performance from the Free-recall phase as a predictor in the LMM (Fig. 5). There was a main effect of identity memory performance on time to first fixation,  $\beta = -0.046$ ,  $SE = 0.012$ ,  $t = -3.88$ , 95 % CI  $[-0.070, -0.023]$  — when participants had previously recalled objects they generally also found them faster. Additionally, we found a significant main effect of total gaze durations on search time,  $\beta = -0.758$ ,  $SE = 0.023$ ,  $t = -3.26$ , 95 % CI  $[-0.121, -0.030]$ . In order to investigate the interaction between active object handling and task relevance we focused our analysis on target objects. Critically, search times in the Find condition were predicted by identity memory performance,  $\beta = -0.084$ ,  $SE = 0.023$ ,  $t = -3.58$ , 95 % CI  $[-0.129, -0.038]$ , whereas this was not true in the Handle condition,  $t < 1$ , demonstrating the invariance of location memory to extant identity representations after object handling. This invariance was on the one hand due to a deprioritization of irrelevant information, as search times for previously recalled irrelevant objects were significantly slower in the Handle (2,283 ms), compared to the Find condition (1,963 ms),  $\beta = -0.076$ ,  $SE = 0.023$ ,  $t = -3.26$ , 95 % CI  $[-0.115, -0.021]$ . On the other hand, previously not-recalled relevant objects were found faster in the Handle (2,238 ms) compared to the Find condition (2,427 ms),  $\beta = 0.054$ ,  $SE = 0.026$ ,  $t = 2.03$ , 95 % CI  $[0.002, 0.105]$ , indicating the prioritization of location information from relevant handled objects beyond the availability of identity memory. At the same time, search times for previously recalled relevant objects did not significantly differ between the Handle and the Find condition,  $t < 1$ .

### Discussion

Ongoing interactions with the external world are an intricate part of natural behavior, and cognitive processes should be investigated in the light of available actions (Clark, 1999; Engel et al., 2013; Gibson, 1979; McGann, 2007; Witt & Riley, 2014; Witt, 2011). Searching for objects in our visual environment can by itself be considered an active exploratory task that constitutes a large portion of our everyday lives. This active behavior can result in memory representations of our



**Fig. 5** Time to first fixation as a function of the experimental conditions and identity memory performance. Search time for objects which were previously recalled in the Free-recall phase is depicted with a black border. Time to first fixation for objects whose identity was not recalled is depicted with gray borders

surroundings that are superior even to explicit memorization (Draschcow et al., 2014; Vö & Wolfe, 2012). In many cases we search for objects that have an immediate relevance to current tasks and as a consequence actively engage with these objects, e.g., picking up the keys you had been looking for. Object recognition speed (Harman et al., 1999; James et al., 2002) and location memory (Tatler et al., 2013; Trewartha et al., 2015) increase after object manipulation. However, there is evidence that we are not necessarily better in recalling/recognizing active compared to passive objects (Brooks et al., 1999; Harman et al., 1999; James et al., 2002) and are not always left with a better spatial representation of our surroundings after active object manipulation (Kirtley & Tatler, 2015). Within a naturalistic, real-world paradigm our study investigated the role of active object handling on identity and location memory for objects. To the best of our knowledge, we are the first to compare memory performance between handled and non-handled task-relevant objects within a real-world environment, as in previous studies the non-manipulated objects were either always irrelevant to the task (Kirtley & Tatler, 2015) or presented on a computer screen (Tatler et al., 2013). We show that: (1) identity memory was not influenced by object handling itself, but relevant objects were recalled better than irrelevant ones; (2) identity memory was highly predictive of location memory, as recalled objects were subsequently found faster (this was, however, only true for passive objects, as location memory for actively handled objects was not predicted by identity memory performance); and (3) relevant objects were found faster than irrelevant ones when actively handled, but this was not true for passive objects. Taken together, our results suggest that active object handling leads to spatial representations independent of object identity memory and facilitates the prioritization of relevant location information, while deprioritizing irrelevant information.



One critical aspect of real-world interactions is that different tasks and interaction types lead to different gaze behavior (Hayhoe & Ballard, 2014; Land & Hayhoe, 2001; Tatler, Hayhoe, Land, & Ballard, 2011). Not only do objects relevant to the current task receive more fixations (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003), but time spent fixating these objects also predicts if that object is remembered better (Hollingworth & Henderson, 2002) or found faster (Hollingworth, 2012). In our study, participants were equipped with mobile eye tracking glasses to provide us with information regarding the gaze behavior connected to the experimental manipulation. Feeding gaze duration information into linear mixed-effects models is a powerful tool, as it allows us to differentiate between effects caused by the experimental conditions alone and effects that can be traced back to mere differences in gaze behavior. In the Packing phase, participants had to either search for or collect objects over the course of 80 trials within the same environment. In this first phase, participants improved in search time across trials. To investigate if the accumulation of gaze durations on objects was predictive of the time to first fixation once that object became a target, we included total incidental fixations into our linear mixed-effects model analysis. This provided us with a nuanced pattern of results. Total incidental fixation durations significantly predicted search time and, more importantly, they were the cause for the search time improvement across trials. Yet, many objects which participants had not fixated before were found faster than objects receiving multiple fixations and participants did not become faster with repeated exposure to the same environment when the covariate was included, suggesting that contextual information of the environment can guide search in many cases (Kit et al., 2014; Vö & Wolfe, 2012). At the same time, total incidental fixation did predict time to first fixation, supporting evidence that memory accumulated over trials does get used in subsequent search (Hollingworth, 2012). We included the total time participants spent fixating each object in our analysis of identity and location memory in order to account for this important part of natural behavior. The effects discussed further were above and beyond fixation duration, suggesting that even though gaze duration is critical for subsequent recollection (Hollingworth & Henderson, 2002), experimental instructions had substantial influence on how information was extracted and retained from fixations above and beyond mere gaze durations (Tatler & Tatler, 2013). Further support for this notion comes from the data presented here, which show that not only relevant information was recalled significantly better than irrelevant information in the Free-recall phase, but also fixation durations did not significantly predict recall memory performance. This is in line with the finding that memory for search targets is better predicted by a non-viewing factor (e.g., whether the participant found the target) than by fixation durations on objects (Vö & Wolfe, 2012; Williams, 2010). Successful

searches are especially effective when the target is embedded in a meaningful surrounding, as participants remember them even better than intentionally memorized ones. However, this seems not to be true for objects presented in isolation (Draschkow et al., 2014).

In our study, active object handling did not influence identity memory for these objects. Even though previous studies have demonstrated a benefit in recognition speed of actively manipulated objects, our finding is in line with previously reported null results concerning the recognition accuracy difference between active and passive objects (e.g., Harman et al., 1999; James et al., 2002). Additionally, actively controlled movements through a virtual environment strengthen spatial representations of same environment, but do not facilitate object recognition and recall of objects compared to passive observation (Brooks et al., 1999). This might, in part, be explained by the active nature of the “Find condition” of our and the Brooks and colleagues study, as merely searching for objects embedded in a meaningful context can be considered active and has been shown to result in superior object identity memories even compared to ones generated during explicit memorization (Draschkow et al., 2014). Together these findings indicate that there might be facilitated processing of actively handled objects, but this does not necessarily result in better memory for these objects in real-world environments in which participants are performing natural tasks.

When handling objects during interactions with our surroundings, it is intuitive to assume that the location of these objects is prioritized by our cognitive apparatus when compared to simply observing them. In fact, active hand movements leave participants with better object location memories than when the hand is passively moved, irrespective of the way memory is inferred – correct responses or subsequent search time (Trewartha et al., 2015). Location memory for actively handled task-relevant objects is superior to memory for objects whose handling was only passively observed (Tatler et al., 2013). However, similar to Brooks et al. (1999), we found no main effect of object handling. Yet, there was a significant interaction between active object handling and task relevance. Critically, relevant objects were found faster than irrelevant ones in the Handle but not in the Find condition. This prioritization of task-relevant information in the Handle condition falls in line with the findings of Tatler and colleagues and demonstrates the close relationship between the constraints of the task and the opportunity for action.

Finally, in our study participants had to recall all object identities they could remember, but then also had to subsequently search for all critical objects in the environment. Participants were not aware of the unannounced memory task, yet a mean of about 50 % of the 80 target objects were recalled, demonstrating strong incidental encoding even when the information is presumably not needed after the completion of the 80 trials (Castelhano & Henderson, 2005; Draschkow

et al., 2014; Hollingworth, 2006; Vö, Schneider, & Matthias, 2008). Nevertheless, no participant could recall all objects, suggesting that participants subsequently searched for objects whose identity representation they had either forgotten or could not access. Our paradigm allowed us to include identity memory performance as a covariate in the linear mixed-effects model of location memory. Location memory has been shown to be strongly related to identity memory (Olson & Marshuetz, 2005), and in fact location memory for recalled objects is very high after incidental memorization during search (Draschkow et al., 2014), suggesting that location memory comes “for free” if you are able to recall an object’s identity. Accounting for identity memory performance when investigating location memory for the same objects allowed us to dissociate between “what” and “where” representations. As expected, objects which were remembered in the Free-recall phase were found significantly faster during the Search phase. Indeed, it is possible that the process of accurately recalling an object in the Free-recall phase might exert a modulatory role on the subsequent search and thus confounds this result. However, this could not explain the striking finding that search time for objects which were previously handled by the participants remained unaffected by previous identity memory performance.

Task-relevant objects whose identity was not remembered were found faster in the Handle compared to the Find condition, indicating that location memory for these items was prioritized. Search times for irrelevant objects whose identity was remembered, however, were increased after active object handling. The pattern of data suggests that location memory for objects is not static and object handling influences relevant and irrelevant information differently. One explanation for the decreased memory performance of handled irrelevant objects might be that objects that were collected no longer contained a single location code because the participants actively moved the objects themselves. This movement could have updated the location of the object in memory, resulting in a fuzzier location representation, and therefore could have made it more difficult to remember its original location. Alternatively, two location codes might have remained in memory for the recalled objects (the original and the new location) with the latter location causing retroactive interference on the former. The fact that previously recalled relevant objects did not significantly suffer from handling might speak against these explanations, yet there was at least a numerical decrease in performance for previously recalled relevant objects. Finally, the act of moving/stowing away a task-irrelevant object might facilitate deprioritization or even induce forgetting of the object’s location information. In other words, participants might have implicitly discarded the information by literally “putting away” the irrelevant object location out of memory. However, more research is needed to investigate those diverging explanations.

To conclude, natural behavior within a realistic environment leaves participants with reliable incidentally generated memory representations of object identities and their locations. “What” information is generally highly predictive of “where” information, yet active object handling can support object location memory beyond a participant’s ability to recall the identity of these objects. The goal-oriented fashion of real-world interactions with the external world leads to the prioritization of relevant location information. Object handling seems to facilitate this process, specifically when it comes to location information of critical objects of which the identity information is missing. This might go in hand with deprioritizing irrelevant location information of objects when their identity information is available.

**Acknowledgments** This work was supported by a Deutsche Forschungsgemeinschaft (DFG) Grant VO 1683/2-1 to MLV. We wish to thank Carrick Williams, and anonymous reviewers for their helpful comments, as well as Sage Boettcher, Daniela Gresch, Maximilian Scheuplein, and Leonie Polzer for valuable help with data collection and analysis.

## References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. doi:10.1016/j.jml.2007.12.005
- Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious Mixed Models. *Journal of Memory and Language*, 27. Methodology. Retrieved from <http://arxiv.org/abs/1506.04967>
- Bates, D. M., Maechler, M., Bolker, B. M., & Walker, S. (2014). lme4: linear mixed-effects models using Eigen and S4. *R package version, 1*, 1–7. Retrieved from <http://cran.r-project.org/package=lme4>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9176952>
- Brooks, B. M., Attree, E. A., Rose, F. D., Clifford, B. R., & Leadbetter, A. G. (1999). The specificity of memory enhancement during interaction with a virtual environment. *Memory (Hove, England)*, 7(1), 65–78. doi:10.1080/741943713
- Castelhano, M. S., & Henderson, J. M. (2005). Incidental visual memory for objects in scenes. *Visual Cognition: Special Issue on Real-World Scene Perception*, 12, 1017–1040.
- Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, 9(3), 6.1–15. doi:10.1167/9.3.6
- Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, 3(9), 345–351. doi:10.1016/S1364-6613(99)01361-3
- Development Core Team, R. (2012). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org>
- Draschkow, D., Wolfe, J. M., & Vö, M. L. H. (2014). Seek and you shall remember: Scene semantics interact with visual search to build better memories. *Journal of Vision*, 14(8), 10. doi:10.1167/14.8.10
- Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where’s the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*, 17(5), 202–209. doi:10.1016/j.tics.2013.03.006

- Findlay, J. M., & Gilchrist, I. D. (2001). In M. Jenkin & L. Harris (Eds.), *Vision and attention*. New York: Springer. doi:10.1007/978-0-387-21591-4
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology*, 9(22), 1315–1318. doi:10.1016/S0960-9822(00)80053-6
- Hayhoe, M. M., & Ballard, D. (2014). Modeling task control of eye movements. *Current Biology: CB*, 24(13), R622–R628. doi:10.1016/j.cub.2014.05.020
- Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, 3(1), 49–63. doi:10.1167/3.1.6
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7(11), 498–504. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14585447>
- Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, 16(4), 219–222. doi:10.1111/j.1467-8721.2007.00507.x
- Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 58–69. doi:10.1037/0278-7393.32.1.58
- Hollingworth, A. (2012). Task specificity and the influence of memory on visual search: Comment on Vö and Wolfe (2012). *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1596–1603. doi:10.1037/a0030237
- Hollingworth, A., & Henderson, J. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 28(1), 113–136.
- James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. A. (2002). “Active” and “passive” learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, 34(3), 383–390. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12395554>
- Kirtley, C., & Tatler, B. W. (2015). Priorities for representation: Task settings and object interaction both influence object memory. *Memory & Cognition*. doi:10.3758/s13421-015-0550-2
- Kit, D., Katz, L., Sullivan, B., Snyder, K., Ballard, D., & Hayhoe, M. (2014). Eye movements, visual search and scene memory, in an immersive virtual environment. *PloS One*, 9(4), e94362. doi:10.1371/journal.pone.0094362
- Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2010). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*, 1, 238. doi:10.3389/fpsyg.2010.00238
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25–26), 3559–3565. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11718795>
- McGann, M. (2007). Enactive theorists do it on purpose: Toward an enactive account of goals and goal-directedness. *Phenomenology and the Cognitive Sciences*, 6(4), 463–483. doi:10.1007/s11097-007-9074-y
- Olson, I. R., & Marshuetz, C. (2005). Remembering “what” brings along “where” in visual working memory. *Perception & Psychophysics*, 67(2), 185–194. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15971683>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9176953>
- Tatler, B. W., Gilchrist, I. D., & Land, M. F. (2005). Visual memory for objects in natural scenes: From fixations to object files. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, 58(5), 931–960. doi:10.1080/02724980443000430
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting saliency. *Journal of Vision*, 11(5), 5. doi:10.1167/11.5.5
- Tatler, B. W., Hirose, Y., Finnegan, S. K., Pievilainen, R., Kirtley, C., & Kennedy, A. (2013). Priorities for selection and representation in natural tasks. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 368(1628), 20130066. doi:10.1098/rstb.2013.0066
- Tatler, B. W., & Tatler, S. L. (2013). The influence of instructions on object memory in a real-world setting. *Journal of Vision*, 13(2), 5. doi:10.1167/13.2.5
- Thomas, L. E., Davoli, C. C., & Brockmole, J. R. (2012). Interacting with objects compresses environmental representations in spatial memory. *Psychonomic Bulletin & Review*, 20(1), 101–107. doi:10.3758/s13423-012-0325-8
- Trevartha, K. M., Case, S., & Flanagan, J. R. (2015). Integrating actions into object location memory: A benefit for active versus passive reaching movements. *Behavioural Brain Research*, 279, 234–239. doi:10.1016/j.bbr.2014.11.043
- Vö, M. L.-H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, 9(3), 24.1–15. doi:10.1167/9.3.24
- Vö, M. L.-H., Schneider, W., & 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.
- Vö, M. L.-H., & Wolfe, J. M. (2012). When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 38(1), 23–41. doi:10.1037/a0024147
- Vö, M. L.-H., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, 1339, 72–81. doi:10.1111/nyas.12667
- Williams, C. C. (2010). Not all visual memories are created equal. *Visual Cognition*, 18(2), 201–228. doi:10.1080/13506280802664482
- Witt, J. K. (2011). Action’s effect on perception. *Current Directions in Psychological Science*, 20(3), 201–206. doi:10.1177/0963721411408770
- Witt, J. K., & Riley, M. A. (2014). Discovering your inner Gibson: Reconciling action-specific and ecological approaches to perception–action. *Psychonomic Bulletin & Review*, 21(6), 1353–1370. doi:10.3758/s13423-014-0623-4