

Recalling what was where when seeing nothing there

Maria Staudte¹ · Gerry T. M. Altmann²

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract So-called “looks-at-nothing” have previously been used to show that recalling *what* also elicits the recall of *where* this was. Here, we present evidence from an eye-tracking study which shows that disrupting looks to “there” does not disrupt recalling what was there, nor do (anticipatory) looks to “there” facilitate recalling what was there. Therefore, our results suggest that recalling *where* does not recall *what*.

Keywords Episodic memory and recall · Eye movements and visual attention · Retrieval cues and memory · Spatial memory

Recent research has shown that people fixate an empty region on a screen if this region had previously been occupied by an object that is now retrieved from memory (Altmann, 2004; Hoover & Richardson, 2008; Johansson, Holsanova, & Holmqvist, 2006; Richardson & Spivey, 2000; Richardson, Altmann, Spivey, & Hoover, 2009). The functional role of these so-called *looks-at-nothing* has been shown to facilitate memory retrieval of both detailed spoken information and visual features of depicted objects

(Hollingworth, 2009; Laeng & Teodorescu, 2002; Laeng, Bloem, D’Ascenzo, & Tommasi, 2014; Scholz, Mehlhorn, & Krems, 2016). In these studies, a reference to an object was typically *given* and visual details of this object were to be encoded and later retrieved by participants. Looks at the (previous) location of this object were shown to help participants recall the object’s features. Further, when the object location was given in addition to the object reference, as in Johansson and Johansson (2014) or Scholz et al. (2016), retrieval was again facilitated compared to when a wrong, incongruent location cue was given. These results suggest that when knowing the identity *and* previous location of an object, looking at this location benefits retrieval of object details. Thus, when the identity of an object is known to the participant, additionally cueing its previous location can help retrieve visual features of this object. But is it equally possible that when *only* the location can be recalled, this facilitates the retrieval of the *identity* of the object which was at that location? That is, does knowing — and looking at — a location help to retrieve what was there?

Previously, Altmann and Kamide (2007) proposed that people *look-at-nothing* because when the episodic trace of an object is re-activated, its location is activated as well, thus, the eyes move. Although the Hebbian account given there might predict that activating its location should also activate the contents of the episodic trace, there is little evidence to support a fully bidirectional account (see Richardson et al., 2009, for a discussion). One obvious reason for a weaker version, e.g. “weighted” bidirectional activation with stronger activation in one direction and weaker activation in the other, could be that the previous location of an object may have hosted several other objects, some more recently even; object identity may be a probabilistically better cue to a specific prior location than prior location is a cue to a specific object identity. Essentially, whereas when I

✉ Maria Staudte
masta@coli.uni-saarland.de

¹ Department of Computational Linguistics, Saarland University, Campus, C7 4, 66123 Saarbrücken, Germany

² CT Institute for the Brain and Cognitive Sciences and Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA

look at an object I necessarily experience a location, when I look at a location, I do not necessarily experience an object. Nonetheless, it is plausible to assume that activating (by looking at) the previous location of a specific and known object can activate further information about this object. But it is unclear whether activating (and looking at) a location alone (in the absence of activated identity information) does also activate the identity of the specific object that was there.

The aim of this paper is to shed light on this issue by presenting evidence indicating that eye-movements to *where* an object had been do not in fact facilitate retrieval of *what* had been there. To enable participants to retrieve the identity of a to-be-recalled-object by themselves, we employed a task in which participants learned sequences of letters, each presented sequentially at a distinct location in a grid (cf. Lange & Engbert, 2013). We then probed either the recognition of a sequence of *locations* at which the letters occurred, regardless of the actual letters in those locations, or the sequence of *letters*, regardless of where they had been. Since anticipation of an upcoming item in such a sequence requires active retrieval from memory, each item incurred a retrieval and a recognition phase (corresponding to anticipating, and then perceiving). This sort of probing further allowed us to distinguish between retrieval of spatial information (*where* something appeared at a particular sequential position) and identity information (*what* that something was). To preview our results: We found that *fixed viewing* at recall influenced recall of location, but not identity information. In *free-viewing*, spontaneous anticipatory eye-movements to correct locations of upcoming items suggested correct retrieval of letter location and identity. However, they did not lead to significantly higher accuracy than spontaneous

eye-movements to other, potentially incorrect locations (or simply staying central on the screen voluntarily).

Specifically, our design was as follows: Participants saw a 5×5 grid which was filled with letters. A circle appeared and highlighted one letter before moving to another, highlighting a sequence of five letters in total, each in a different location (see Fig. 1). The task was to memorize both the LOCATION and the IDENTITY of these highlighted letters. In contrast to the task and procedure employed by Lange and Engbert (2013) and similar to Johansson and Johansson's study, our participants did not know whether they would have to recall identity or location of the (sequence of) letters until recall started. A blank grid was then presented and either a circle moved across a sequence of blank squares, representing object locations without identity, or letters were presented verbally, reflecting object identity but not location. In principle, participants could anticipate and thereby *actively retrieve* an item in the serial presentation before its location or identity was actually given. The retrieved information could then be compared with the subsequently presented location or identity (c.f. a recognition phase). We were therefore able to distinguish “congruent” trials with correct anticipatory eye-movements to the location of an upcoming item from “incongruent” trials with other non-anticipatory eye-movements. We further manipulated eye-movements by either allowing free-viewing or by using a fixation cross to constrain eye-movements during the entire test (i.e., retrieval and recognition) phase. Participants had to press a button at the end of each sequence to indicate whether or not the test sequence was correct, i.e., identical with the encoded sequence. Eye-movements were recorded and analysed throughout the whole test phase. The two tasks we employed generate distinct predictions:

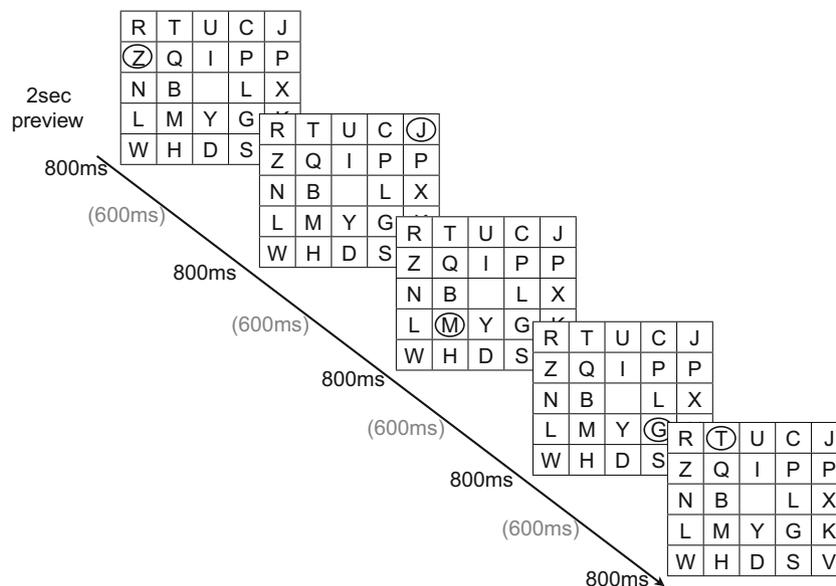


Fig. 1 Encoding a sequence. Here the sample letter sequence “Z-J-M-G-T” is shown

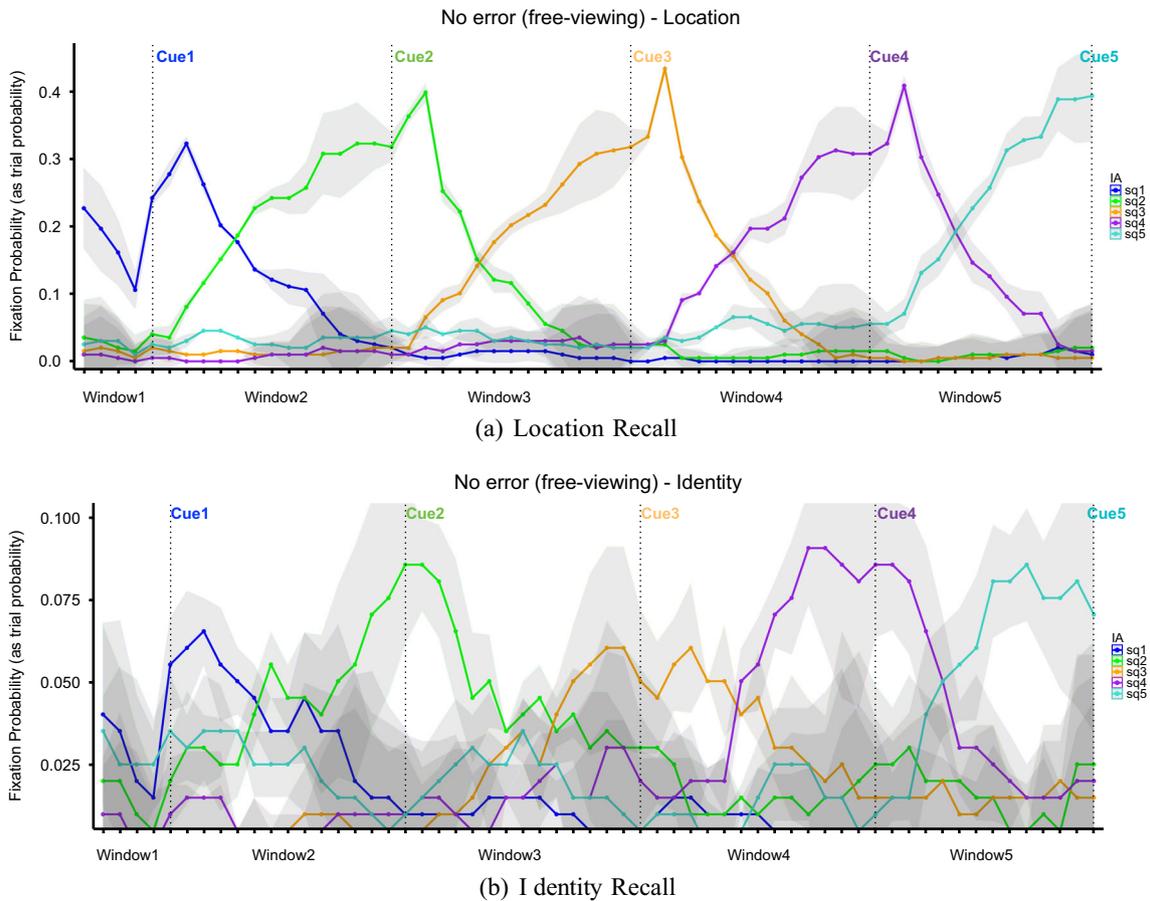


Fig. 2 Time graph (in 100ms bins) which plots fixation probability to the squares involved in the test sequence of (a) letter locations, and (b) letter identity. The lines representing mean probabilities are accompanied by confidence interval areas shaded in grey. Vertical lines in the

graphs symbolize the onset of a cue (to object identity or location) and therefore the border to the next time window. Note the different scales for Location and Identity recall

- **Location Recall:** The location recall task required participants to correctly match the test sequence of locations against the previously presented sequence (which contained both location and identify information). It

was included to check that location was indeed encoded and successfully retrieved. We predicted anticipatory eye-movements to correct locations, reflecting anticipation of the imminent appearance of the circle in the

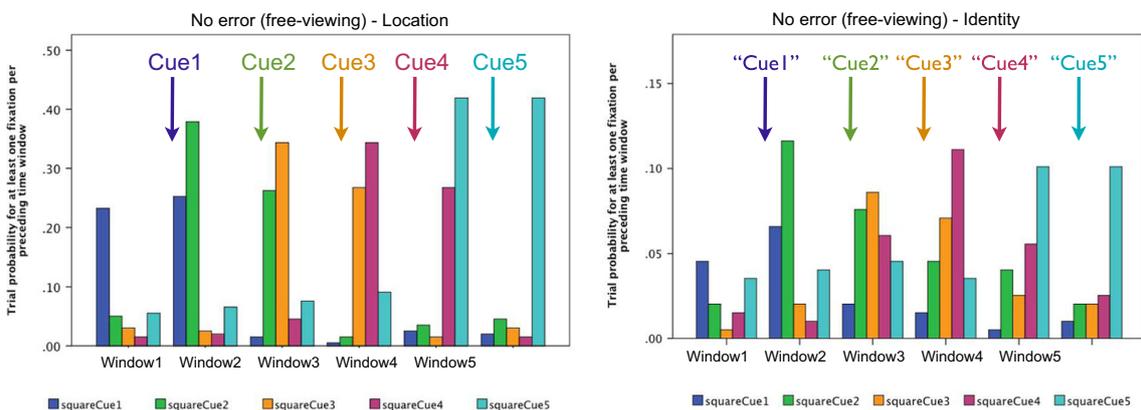


Fig. 3 Trials with no error and free-viewing in the test phase: Trial probability for at least one fixation to a relevant square in a given time window. Time windows span the 1400ms between cue onsets

Table 1 Fixation probabilities for critical square compared to average using binomial test

	Window1	Window2	Window3	Window4	Window5
Location, noErr					
CritSquare vs AvgSquare	0.05, 0.02	0.12, 0.04	0.08, 0.05	0.12, 0.06	0.11, 0.04
p-Value	< .05	< .001	<.05	< .01	< .001
Identity, noErr					
CritSquare vs AvgSquare	0.25, 0.08	0.41, 0.16	0.37, 0.16	0.37, 0.15	0.45, 0.16
p-Value	< .001	< .001	<.001	< .001	< .001

next location (i.e., correct retrieval), as well as eye-movements to these locations once the circle appeared, reflecting oculomotor capture. We also predicted that fixed viewing would lead to lower accuracy than free-viewing if participants were re-enacting their eye-movements from the encoding phase, as in Bochynska and Laeng (2015). Lastly, we predicted that anticipatory eye-movements to the correct location (congruent with the upcoming stimulus and reflecting correct active retrieval) would be accompanied by higher accuracy than when no such eye-movements were observed.

- **Identity Recall:** The identity recall task required participants to listen to a stream of five letters and respond whether the sequence matched the previously presented one. If eye-movements to letter locations in this task are indeed helpful for retrieval and/or recognition, we should observe lower accuracy when the participant fixates centrally compared to free-viewing. Moreover, anticipatory eye-movements to the location associated with the next letter in the sequence would indicate that the participant knew the letter's location (and possibly its identity) *before* the item appeared. If participants had successfully retrieved location *and* identity, or had not retrieved identity until it was re-activated by looking at that location, we would predict (a) higher accuracy for trials with anticipatory eye-movements ("congruent" trials) than for trials without, and (b) lower accuracy in the fixed viewing condition (where anticipatory eye-movements cannot be launched).

Method

The experiment was conducted as a 2×2 within-subjects design. The first factor (Task) manipulated the type of information that had to be retrieved/recognized, that is, whether letter location or letter identity was requested. The second factor (FixCross) manipulated whether participants were able to freely move their eyes during recall or whether they had to fixate centrally. Thirty-four native speakers

of English, all but one enrolled at the University of York as undergraduate students of Psychology, took part in this study and received course credits for participation. Two participants were excluded from analyses since they were run on the wrong script resulting in a total of thirty-two participants. All reported normal or corrected-to-normal vision.

Materials

For items, we randomly sampled five letters from the four main quadrants (2×2) in the grid, sparing the centre lines, in a way such that the first four letters were all taken from different quadrants (see Fig. 1). For fillers, we sampled from all grid cells. Letter sequences were designed not to contain actual words or known abbreviations. During location recall, the circle was presented on a grid of empty cells while letter (identity) recall consisted of utterances of each letter produced by a text-to-speech system.

We created a total of 48 items and 24 fillers, distributed equally across six different blocks. Each block showed a grid with a different arrangement of letters. In three out of six blocks, participants had to fixate a fixation cross during recall. Whether letter identity or location was probed varied equally within each block. Further, half of the item sequences presented during recall contained an error. That is, a letter taken from the close spatial vicinity of the original letter replaced it. This only affected letters with a serial position of 3, 4 or 5 in the sequence. The reasons for not including errors in serial position 1 or 2 was that it could have rendered those sequences uninformative for further eye-movement analyses if participants stopped paying attention after detecting an (early) error.

Table 2 Mean accuracy for responses in each condition

	FREE-VIEWING	FIXATION CROSS
LOCATION	77.9	72.9
IDENTITY	82.5	82.7

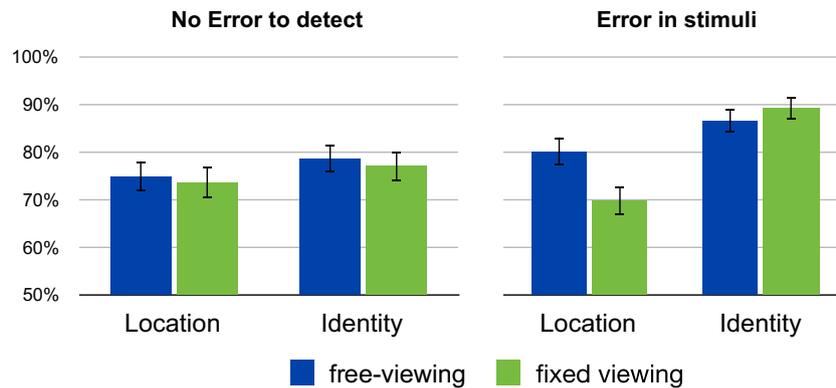


Fig. 4 Accuracy of responses in % with standard error bars, for sequences that contained an error and those that did not. Error detection for object location was more difficult, i.e., elicited the lowest accuracy, when participants were requested to keep looking at the fixation cross

Task and procedure

Each session started with the 9-point calibration of the EYELINK II eye-tracker sampling at 500Hz. Calibration was repeated between blocks. Participants were instructed to memorise sequences of letters and their locations and that they were then asked to validate a repetition of only the locations or only the letters as correct or incorrect. The presentation of the original sequence started 2000ms after the grid filled with letters was displayed. Each circle appeared for 800ms, with an inter-stimulus-interval (ISI) of 600ms (cf. Fig. 1). After the encoding phase, the word “Letters” or “Locations” appeared on the screen, indicating which type of information would be requested and presented in the test phase. Then, either circles or spoken letters were provided with the same duration and ISI as during encoding. After the whole sequence was presented, responses were elicited by displaying the text “Sequence correct?”. The experiment lasted approximately 45 minutes overall.

Analysis

We segmented the test sequence into five adjacent time windows. “Window1” contains the initial 600 ms before presentation of cue one (“Z” in the example in Fig. 1). “Window2” includes the presentation of the “Z” and the ISI, with 1400 ms duration in total. Three more time windows are coded analogously to Window2 (corresponding to “J”, “M”, and “G” from Fig. 1). We also coded six regions of interest: one for each square that contained a critical letter and one for the square that contained the alternative letter composing the error. For instance, the square containing the “J” in the sample sequence “Z-J-M-G-T” would be the region of interest “square 2”. The time window before its appearance (Window2, “Z”+ISI) would be associated with anticipatory eye-movements to “J” (retrieval phase) whereas Window3 (“J”+ISI) would be associated with eye-movements

concurrent with it (c.f. a recognition phase), as well as with anticipatory eye-movements to the next letter (“M”).

Inferential statistics were carried out using mixed-effects models from the lme4 package in R (Baayen, Davidson, & Bates, 2008). Specifically, we used logistic regression for modeling accuracy of participants’ button presses. Fixed effects and interactions were determined through model selection. Random intercepts and slopes were included in models when they were justified by the design (Barr, Levy, Scheepers, & Tily, 2013) and then reduced until the models converged. We further used the binomial test to compare trial counts with and without fixations between squares in order to determine whether the location of the upcoming stimulus was anticipated or not.

Results

Eye-movements

We plotted fixations in each region of interest in 100ms-bins across each test sequence. Figure 2a) shows anticipatory eye-movements as well as oculomotor capture during location recall: Rising fixation proportions to e.g. square 2 before the respective cue appears (Window2), and immediately after its appearance (Window3) before declining again.

Table 3 Model fitted to accuracy data overall in erroneous recall sequences

Predictor	Coeff.	SE	Wald z	p-Value
(Intercept)	2.24	0.319	7.03	<.001
Task-position	-0.53	0.300	-1.77	=.076
FixCross-withFixCross	0.35	0.378	0.92	=.352
position-withFixCross	-0.93	0.434	-2.16	<.05

*Model : Accuracy ~ Task * FixCross + (0 + FixCross|subject)*

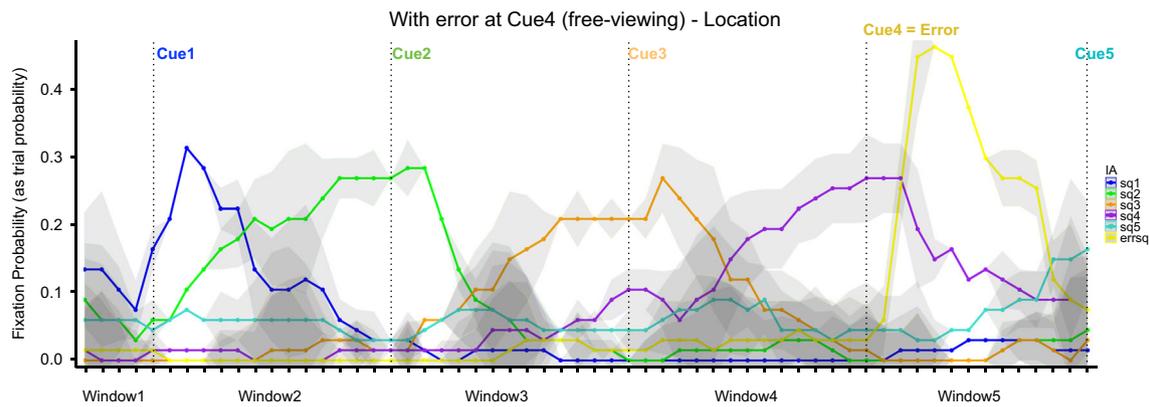


Fig. 5 Fixation time line for sequences that contained an error at position 4, for location recall. Anticipatory eye-movements to the location of cue4 are visible in Window 4, and concurrent eye-movements to the location of the actually appearing error-cue are visible in Window5

Figure 2b) shows that anticipatory eye-movements during letter recall, which reflect *active retrieval* of letter *locations*, also occur during the retrieval and recognition of letter *identity* — even though to a lesser extent than in location retrieval.

For inferential statistics, we calculated the probability for a trial to have at least one fixation in a given time window and square. Figure 3 shows the resulting probabilities for the same conditions as plotted in Fig. 2.

In each time window, fixation probabilities for each critical square were compared with the average fixation probability of the four other squares (including the currently cued one) in this sequence by means of a binomial test.¹ The results of this test for both tasks and for each time window are given in Table 1. In sum, the results show that the upcoming letter always elicited significantly more (anticipatory) eye-movements to the correct square than to the other squares involved in that same sequence.

Behavioral data

We collected a total of 1521 button presses of which 319 (20.9 %) were incorrect responses. Their distribution is shown in Table 2. Generally the task was considered difficult but manageable, as the accuracy rates show. In sum, participants performed worse in validating the location of a stimulus compared to its identity ($\chi^2(1) = 12.91, p < .001$). The fixation cross manipulation did not affect model fit ($\chi^2(1) = 0.75, p = .383$) nor was there an interaction with Task (free-viewing vs. fixed viewing: $\chi^2(2) = 1.30, p = .253$), contrary to what the numeric difference in the location task might suggest.

¹The example for the cell which contained the upcoming item 2 in Window2 required the following syntax in R: `binom.test(22, 189, 0.04, alternative = "greater")` with the outcome of $p < 0.001$.

We then split the dataset into test sequences that contained an error and those that did not in order to examine whether error detection was overall similarly difficult across conditions (Fig. 4). Model comparison on the erroneous test sequences revealed a significant interaction between Task and FixCross ($\chi^2(1) = 4.56, p < .05$) which was carried by the significant difference between the free-viewing and fixed viewing in the location task (Table 3).

Next, we considered accuracy as a function of congruence between anticipatory eye-movements (to the retrieved letter's location) and the actual (recognized) letter in those error sequences. Congruent anticipatory eye-movements to a letter's location indicate correct (location) retrieval, which is the pre-requisite for detecting the error (in location recognition). Thus, these cases are particularly suitable for examining whether or not correct active (anticipatory) retrieval of an object's location accompanies increased accuracy in location recall, and/or facilitates retrieval/recognition of this object's identity in the identity task. In order to examine the benefit of such anticipatory eye-movements for (error) recognition, we considered only trials with a test sequence that contained an error (see e.g. Fig. 5). Further, we split these trials according to whether they contained a correct anticipatory eye-movement to the upcoming letter. That is, given the encoded sequence "Z-J-M-G-T" and an error in position 4, such that "Z-J-M-S-T" actually appeared, we examined whether participants inspected the "G"-location in Window4, immediately before "S" came up, and whether this affected response accuracy (was the sequence more likely to be recognized as wrong if that correct anticipatory eye movement had been made?).

From all error sequences in the location condition, we extracted 76 trials that contained an inspection to the correct letter location in the time window right before the (erroneous) letter appeared and 108 trials that contained inspections to other locations on the screen (6 trials with

no fixation at all). 81.6 % of the trials with a correct anticipatory inspection and 80.5 % of the trials without anticipatory looks were correctly responded to. A generalised linear model with a binomial link function fitted to these accuracies, with “inspection to the correct square before the error” as predictor, shows that this difference was not significant (Coeff.= 0.06, $SE = 0.38$, Wald $Z = 0.17$, $p = 0.86$).

Similarly for the identity condition, we extracted 25 trials that contained an inspection to the correct item location and 160 to other locations (6 trials with no fixation at all). 92 % of the trials with a correct anticipatory inspection and 85 % of the other trials were correctly validated. Again, a generalized linear model shows that this difference was not significant (Coeff.= 0.70, $SE = 0.76$, Wald $Z = 0.91$, $p = 0.35$). While the number of instances for the correctly anticipated items is rather low, the percentage of accurate button presses is nevertheless high in both tasks and conditions.

Discussion and conclusion

Overall, both location and identity of letters were encoded and remembered with high accuracy (72 %-83 %) which was still far from ceiling. We also observed that fixed viewing during recall of a location sequence decreased accuracy, but only when there was an error to be detected. Crucially, having to fixate the fixation cross did not lower accuracy in identity recall significantly. Further, spontaneous anticipatory eye-movements to correct locations of a letter neither enhanced accuracy of location recall *nor* accuracy of identity recall. That is, even though participants remembered letter location well, there were only few eye-movements to it during identity retrieval or recognition *and* these eye-movements were not beneficial for performance.

One might argue a) that location and identity were never encoded together and b) that identity retrieval was easy enough to not require any spatial information. However, the former seems implausible since this was effectively a dual task in which both location and identity always had to be encoded. As suggested by Morey and Cowan (2004), encoding and holding independent verbal and visual information in memory appears to consume shared working memory capacity, with errors in one modality affecting accuracy in the other. This indicates that representations are not stored fully independently. For verbal (in our case identity) and visual (here location) information, which were not independent in our study but were associated with each other, joint representation and storage is even more likely. On the contrary, it is difficult to imagine how identity information could be encoded fully independently of location

information; how else could one retrieve location from identity (the basis for language-mediated eye movements; Altmann & Kamide, 2007)? Regarding the ease of the task, the accuracy of identity recall was in a range that leaves room for improvement; it thus seems unlikely that participants would disregard cues that would facilitate task performance.

Further, we observed only little detrimental influence of the fixation cross on (correct) location recall, in contrast to what we had predicted. We propose that the absence of a clear effect in this particular task may be explained by participants exploiting the benefit of a steady frame of reference for encoding and retrieving spatial patterns by suppressing saccades to relevant locations (as suggested by Lange & Engbert, 2013, for encoding). But more investigation is necessary to establish the link between encoding and decoding strategies along the lines of Bochynska and Laeng (2015).

Finally, our results point to a somewhat weaker bi-directional Hebbian account of activation in episodic memory (Altmann & Kamide, 2007). In particular, we showed that there is no location-cued episodic retrieval in a paradigm where identity and location are truly unrelated and arbitrary. Only when the object identity is known might the activation of its location activate other parts of the object’s trace (e.g., other visual features that were not initially activated). Of course, the situation may be different when objects are in their prototypical locations, i.e., cups on a table, birds in the sky etc. Here, knowing the location of an object may indeed facilitate the retrieval of which specific object had been there. Further research is required to explore these possibilities.

To conclude, we have presented an experimental design which allowed us to investigate whether object recall is facilitated by cueing the ‘semantically unrelated’ location of the to-be-recalled object in a clear and consistent manner. The results of this experiment suggest that there is no location-cued episodic retrieval in paradigms where identity and location are unrelated. That is, there is no advantage of retrieving and, crucially, moving ones eyes to a location for the subsequent retrieval of identity information.

Acknowledgments The research reported in this paper was supported by the “Multimodal Computing and Interaction” Cluster of Excellence at Saarland University and by ESRC grant RES-062-23-2749 awarded to the second author.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Altmann, G.T.M. (2004). Language-mediated eye movements in the absence of a visual world: The 'blank screen paradigm. *Cognition*, 93(2), B79–B87.
- Altmann, G.T.M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57(4), 502–518.
- Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Barr, D.J., Levy, R., Scheepers, C., & Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bochynska, A., & Laeng, B. (2015). Tracking down the path of memory: Eye scanpaths facilitate retrieval of visuospatial information. *Cognitive Processing*, 16(Suppl 1), S159–S163.
- 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(1-2), 273–291.
- Hoover, M.A., & Richardson, D.C. (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition*, 108(2), 533–542.
- Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science*, 30(6), 1053–1079.
- Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, 25(1), 236–242.
- Laeng, B., Bloem, I.M., D'Ascenzo, S., & Tommasi, L. (2014). Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, 131(2), 263–283.
- Laeng, B., & Teodorescu, D.S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, 26(2), 207–231.
- Lange, E.B., & Engbert, R. (2013). Differentiating between verbal and spatial encoding using eye-movement recordings. *The Quarterly Journal of Experimental Psychology*, 66(9), 1840–1857.
- Morey, C.C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*, 11(2), 296–301.
- Richardson, D.C., Altmann, G., Spivey, M.J., & Hoover, M.A. (2009). Much ado about eye movements to nothing: A response to Ferreira et al.: Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, 13(6), 235–236.
- Richardson, D.C., & Spivey, M.J. (2000). Representation, space and hollywood squares: looking at things that aren't there anymore. *Cognition*, 76(3), 269–295.
- Scholz, A., Mehlhorn, K., & Krems, J.F. (2016). Listen up, eye movements play a role in verbal memory retrieval. *Psychological Research*, 1–10.