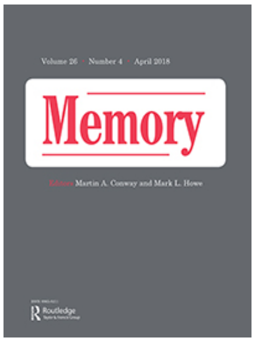

Memory

ISSN: 0965-8211 (Print) 1464-0686 (Online) Journal homepage: www.tandfonline.com/journals/pmem20

Distinct eye movements for different cognitive processes as expressed in the face recognition task

Ohad Parag & Eli Vakil

To cite this article: Ohad Parag & Eli Vakil (2018) Distinct eye movements for different cognitive processes as expressed in the face recognition task, Memory, 26:4, 524-534, DOI: 10.1080/09658211.2017.1381265

To link to this article: https://doi.org/10.1080/09658211.2017.1381265

Published online: 27 Sep 2017.

MEMORY, 2018
VOL. 26, NO. 4, 524–534
https://doi.org/10.1080/09658211.2017.1381265

Distinct eye movements for different cognitive processes as expressed in the face recognition task

Ohad Parag and Eli Vakil

Department of Psychology and Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, Israel

ABSTRACT
Contemporary research literature indicates that eye movements during the learning and testing phases can predict and affect future recognition processes. Nevertheless, only partial information exists regarding eye movements in the various components of recognition processes: Hits, Correct rejections, Misses and False Alarms (FA). In an attempt to address this issue, participants in this study viewed human faces in a yes/no recognition memory paradigm. They were divided into two groups – one group that carried out the testing phase immediately after the learning phase ($n = 30$) and another group with a 15-minute delay between phases ($n = 28$). The results showed that the Immediate group had a lower FA rate than the Delay group, and that no Hit rate differences were observed between the two groups. Eye movements differed between the recognition processes in the learning and the testing phases, and this pattern interacted with the group type. Hence, eye movement measures seem to track memory accuracy during both learning and testing phases and this pattern also interacts with the length of delay between learning and testing. This pattern of results suggests that eye movements are indicative of present and future recognition processes.

ARTICLE HISTORY
Received 3 May 2016
Accepted 14 September 2017

KEYWORDS
Eye movements; recognition; face processing

The theoretical question of interest addressed is this article is whether eye movements play a role in recognition processes, and if this relationship differs between the four specific recognition conditions that will be presented below. Human recognition studies are usually composed of a learning phase in which stimuli are presented, and a testing phase in which stimuli that were previously shown (old) and stimuli that were not previously shown (new) during the learning phase are presented. In most cases, the participant's task is to decide whether or not each stimulus was presented in the learning phase (Tanaka & Farah, 1993). Consequently, there are four possible memory outcomes during an old–new recognition test, i.e. Hits, Correct Rejections (CRs), Misses and False Alarms (FAs) (Snodgrass & Corwin, 1988). These definitions are borrowed from the Signal Detection Theory which attempts to explain participants' decisions regarding the presence of a stimulus (Macmillan, 2002).

Performance in the four recognition conditions (Hit, CR, Miss and FA) can be summarised using the Hit and FA rates of the two independent conditions (Abdi, 2007). As a result, most behavioural recognition studies only report participant Hit and FA rates (Stanislaw & Todorov, 1999). Since the other components (Miss and CR) can be generated from the Hit and FA rates, the former are considered redundant in terms of behavioural measures. However, this does not necessarily mean that the generated measures reflect the same cognitive processes. As evidence to this claim, neuroimaging studies have shown that Miss and CR involve qualitative brain activity that differs from Hits and FAs (Mecklinger, 2000; Osipova et al., 2006). Therefore, it is important to test each of the four recognition conditions.

During the testing phase, increased activity was seen in primary visual areas, as measured using functional magnetic resonance imaging (fMRI) when stimuli appeared during the learning phase, regardless of whether the response was correct or not (Slotnick & Schacter, 2004). While viewing studied faces prior to explicit identification, it is possible to distinguish between studied faces and faces that were incorrectly endorsed as old (FA) (Hannula, Baym, Warren, & Cohen, 2012). These studies also highlight the fact that behavioural measures provide only partial information about recognition processes, since the primary visual areas in the first study and the eyes in the second study "remembered" the stimuli, but this "memory" was not necessarily reflected in the participants' behaviour. Hence, it is important to differentiate between the four recognition processes without relying solely on participants' answers.

One method for doing so is to use eye movement recordings which have a millisecond resolution and can indicate the exact location of the participant's eyes

CONTACT Eli Vakil vakile@mail.biu.ac.il Department of Psychology and Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan 52900 Israel

within an image. Eye movement patterns can be described using various indicators related to fixation – a brief period of time in which the eyes barely move (Hannula, Althoff, et al., 2010). Eye movements were found to predict recognition processes independently of the specific type of task given – whether participants were told to simply look at the picture or to identify the manipulation made to the picture between phases (Ryan, Althoff, Whitlow, & Cohen, 2000). This automaticity of eye movements is also manifested in the "Repetition effect" stimulus that is presented during the learning phase. In the testing phase, it receives longer first and overall fixations and fewer fixations then a new stimulus, regardless of the participant's answer (Chanon & Hopfinger, 2008; Hannula, Althoff, et al., 2010; Ryan, Hannula, & Cohen, 2007). In addition, eye movements were found to affect future recognition. A significant positive correlation was found between number of fixations on faces in the learning phase and recognition performance during the testing phase (Heisz, Pottruff, & Shore, 2013; Kafkas & Montaldi, 2011). However, in this study there was no separate analysis for Hit and FA trials, which means that the recognition performance score was calculated by subtracting participants' FA rate from their Hit rate.

To the best of our knowledge, no prior study tested eye movements separately for the various components of recognition: Hit, CR, Miss and FA. Most studies focused on the relationship between eye movements and accuracy (based on formulas containing Hit and FA rates) (Heisz et al., 2013; Kafkas & Montaldi, 2011) or did not analyse errors at all (Sharot, Davidson, Carson, & Phelps, 2008; Stacey, Walker, & Underwood, 2005). Hence, the purpose and novelty of this study is to analyse eye movements in the different recognition conditions: Hit and Miss responses during the learning phase (defined by the participant's response to the same face during the testing phase) (Osipova et al., 2006); and Hit, CR, Miss and FA trials during the testing phase. Differences between eye movements in the Hit, CR, Miss and FA conditions will reveal the types of eye movement patterns that can be associated with better (or worse) accuracy and with different kinds of decisions (yes or no) made by the participants. Thus, eye movement recordings will increase our understanding of the reason for the occurrence of each of the four different recognition conditions that are created by the combination of the 2×2 decision type (yes or no) and correctness type (correct or incorrect) matrix.

It is important to record eye movements in both the learning and testing phases since many theories and models of human recognition processes have attempted to explain the unique contribution of encoding (learning phase) and retrieval (testing phase) in memory processes (Tulving & Thomson, 1973).

Time interval affects recognition processes (Schacter, 1999). However, there are contradictory findings regarding the effect of the time lapse between encoding and retrieval on recognition accuracy performance. In some studies, this had no effect (Laughery, Fessler, Lenorovitz, & Yoblick, 1974; MacLin, MacLin, & Malpass, 2001), while in other studies, increasing the time delay resulted in decreased recognition accuracy (Koen, Aly, Wang, & Yonelinas, 2013; Mitchell, Brown, & Murphy, 1990). Differentiating between the effect of time delays on Hit rates and FA rates does not clarify this matter. For instance, though Hit rates decreased in response to an increase in time delay, this had no effect on FA rates (Shepherd, Gibling, & Ellis, 1991). In another study, Hit rates decreased and FA rates increased as the time delay increased (Thapar & McDermott, 2001). Thus, it is important to examine the effect of time on recognition processes without relying solely on participants' answers, but also based on eye movement measures, as done in our study.

In conclusion, since it is important to test each of the four recognition conditions (Hit, CR, Miss and FA) without relying on participants' answers, and in order to analyse each trial separately, we used eye movement recordings which were proven to be indicative of memory processes and to have a high temporal resolution (in milliseconds). We used pictures of human faces because of the standardisation made in prior studies that used them as stimuli (Barton, Radcliffe, Cherkasova, & Edelman, 2007; Heisz et al., 2013). We aimed to create experimental conditions that would produce a high percentage of errors (Miss and FA) in order to accumulate sufficient data for error analysis, but without making it too difficult as to compromise performance by causing participants to guess their answers. In order to do so, the number of images studied was relatively large and the duration of the appearance of the faces in the learning phase was relatively short, leading to optimal behavioural performance for error analysis. These experimental conditions were based on data collected from other similar studies with a percentage of errors that was high, but still lower than guessing rates (Heisz et al., 2013; Kafkas & Montaldi, 2011; Smith & Squire, 2008). In addition, the faces were relatively similar to each other (all Caucasian) with no distinct features or unusual expressions, and the distractors selected resembled the targets (in age and gender).

Based on the fixation measures that were found to be related to later overt recognition (Hannula & Ranganath, 2009) and indicative of the repetition effect (Chanon & Hopfinger, 2008; Hannula, Althoff, et al., 2010; Ryan et al., 2007) and of recognition accuracy (Heisz et al., 2013; Kafkas & Montaldi, 2011), we used the following eye movement measures: dwelling time (explained below), fixation rate, first and overall fixation duration and latency to first fixation. We used various eye movement measures, since each is assumed to be related to a different cognitive process. Lai et al. (2013) noted that the number of fixations is often used to show attention distribution. Chanon and Hopfinger (2008) noted that

> … the time before the first fixation on the item, serves as a measure of the attraction, or guidance, of attention. The duration of the first set of fixations is a measure of how long attention dwells on the object when the participant first finds it. The

> total fixation duration during the viewing period indicates how much attention the item receives over the whole scene-viewing period. Finally, the average duration of each fixation measures how long attention is held each time it falls upon the object. (p. 330)

Participants were divided into two groups – one that performed the testing phase immediately after the learning phase, and another that was given a 15-minute break between phases. Despite the unclear effect of time delay on recognition accuracy, delays of less than 15 minutes were shown to be enough to decrease recognition accuracy, when this effect occurred (Colbert & McBride, 2007).

Similarly, studies found that when the distractor faces were similar to the target faces, there was no difference in hit rates between the immediate and delayed tests, even after a seven-day interval. Nevertheless, FA rates increased substantially in the delayed test compared to the immediate test (Chance & Goldstein, 1987). Hence, we predicted that the Delay group would have higher FA rates than the Immediate group, with no difference between Hit rates in both groups. This would indicate that overall, the Immediate group was more accurate than the Delay group.

We defined accuracy by using Corrected Hit (CH) scores, measured by the percentage of Hit rate minus FA rate (Roediger & McDermott, 1994). We decided to apply additional methods for calculating recognition accuracy in order to overcome participants' response bias, if needed, depending on Hit and FA rates (Harrington, 2006).

Instead of referring to the four different recognition conditions (Hit, CR, Miss and FA) in the testing phase as one independent variable with four levels, we divided it to two independent variables with two levels in each: 1. Answer Type – "Yes" (Hit and FA) vs.' No' (CR and Miss); 2. Correctness type – Accurate (Hit and CR) vs. Error (FA and Miss). The reason for doing so is that since this is a preliminary study, it is important to guarantee our full comprehension of the source of the different eye movement patterns in the different recognition conditions, if such differences are found. As discussed above, eye movements provide an index of memory that can be dissociated from an explicit memory response (Hannula et al., 2012). However, eye movements are also related to decision making (Orquin & Loose, 2013). Hence, we wanted to examine the two processes – Correctness and Answer (or decision making), independently.

Since this is a preliminary study, we did not differentiate between the learning and testing phases when predicting the effects of expected correctness in eye movements. However, existing literature clearly and consistently reports unequal patterns of results between the two phases. For instance, a difference in brain activity between Hit and Miss trials was found during the learning phase but not during the testing phase (Osipova et al., 2006). Based on the relationship between eye movements and recognition processes described above (Hannula & Ranganath, 2009; Hannula et al., 2012; Heisz et al., 2013; Kafkas & Montaldi, 2011), we predicted a main effect for correctness during the learning and testing phases, so that there would be different eye movements between Hit and Miss trials in the learning phase and different eye movements between accurate responses (Hit and CR) and errors (Miss and FA) in the testing phase. Consistent with previous findings indicating the lack of a relationship between explicit response and eye movements (Hannula et al., 2012), we predicted that no significant Answer main effect would be found. Nevertheless, (and based on the relationship between eye movements and decision making as described above (Orquin & Loose, 2013)), we predicted an Answer by Correctness interaction.

Since time delay has an effect on the recognition processes (Schacter, 1999) and eye movements are indicative of recognition processes (Hannula, Althoff, et al., 2010; Hannula, Ranganath, et al., 2010), the retention time was expected to have a differential effect on eye movements. This was expected regardless of the behavioural outcomes, since the latter are not always in line with implicit non-behavioural processes (Slotnick & Schacter, 2004). Hence, we predicted that the Immediate group would display a stronger "repetition effect" than the Delay group, as reflected by fewer fixations and longer first and overall fixation durations (Hannula, Althoff, et al., 2010).

Method

Participants

The study was approved, as required, by the ethics committee of Bar-Ilan University. Participants were students at Bar-Ilan University who were randomly assigned to one of two groups. In the Immediate group, there were 30 participants (6 males) with an average age of 24.7 ($SD = 5.92$). In the Delay group there were 28 participants (7 males) with an average age of 24.64 ($SD = 5.49$). The groups did not differ significantly in age, $t(56) = 0.04$, $p = .97$.

Computer and software – stimuli were displayed on a computer with a 15.6 inch PC screen, using the E- PRIME 2.0 software that records the temporal parameter presentations of stimulations, and schedules the appearance of the stimuli with the computer recorded eye movements. Eye movements were recorded using a Senso Motoric Instrument (SMI) RED-M remote eye-tracker. This system allows free-head movements with a sampling rate of 250 Hz, high accuracy of 0.5 °. A 9-point calibration cycle at the beginning of the experiment provided a spatial resolution of 0.1 °. A camera with an infrared source was located at the front of the laptop screen, below the participant's eye level, 60 centimetres from the participant.

Stimuli: colour facial photos of Caucasian adults with neutral facial expressions and with no unusual features, were marked by a blue circle. The learning phase consisted of 60 faces (30 females and 30 males), and the testing phase included 120 faces (60 females and 60 males).

The pictures were taken with permission from the XM2VTS database (Messer, Matas, Kittler, Luettin, & Maitre, 1999).

Tasks and procedure

Participants were tested individually. Upon arrival, they were informed that they would take part in a facial recognition experiment and were instructed to observe the faces that appear on the screen and to try to remember them for a future recognition test (the same instructions appeared later on the computer screen). The instructions given to both groups were identical. Participants were informed that their eye movements would be monitored. The tester than performed an eye calibration check in which participants were instructed to look at a dot that moved across the screen.

Learning phase – in this phase, 60 faces appeared on the screen in random order. Each face appeared for four seconds with an inter-stimulus interval (ISI) of one second between faces. The faces appeared in the centre of the screen, and each appeared only once during the entire phase.

Testing phase – This phase began either immediately (Immediate group) or after 15 minutes (Delay group), in which participants were free to do as they pleased. At the beginning of this phase, another eye calibration check was performed. Throughout this phase, 120 faces appeared successively in the centre of the screen, and participants were instructed to press a button on the keyboard indicating whether it is an old (P) or a new face (Q), respectively. Unlike the learning phase, the face remained on the screen until the participant responded. All faces that appeared during the learning phase (i.e. 60) also appeared in the testing phase, in addition to 60 new faces.

Data analysis

Of the initial 68 participants that participated in the experiment, we removed 10 participants. Hence, the analysis was based on 58 participants. Four participants were eliminated from the result analyses due to unusual Corrected Hits (CH) scores (Hit - FA). Two participants had a CH score of two *SD* or higher than the CH mean of their group. Two participants had a CH score that was more than 2 *SD* below the CH mean of their group. We used especially strict cut-off points in the behavioural analysis (compared to the eye movements analysis) in order to make every possible effort to ensure that the participants that remained in the analysis did not guess their responses (those with a CH score that was more than 2 *SD* below the CH mean of their group), nor did they encounter our stimuli in a previous study, thus improving their performance (those who had a CH score of two *SD* or higher than the CH mean of their group).

Six participants were eliminated from the result analyses due to problems related to their eye movements: (a) one participant due to eye movement recording problems; (b) two participants because only part of the trials performed in the learning phase were recorded; (c) one participant due to extremely unusual eye movement results in the testing phase (for example: fixation rate mean that was more than three *SD* less than the group mean, and a fixation duration mean that was more than three *SD*s greater than the group mean); and (d) two participants because of inaccurate time gaps between the learning and testing phases.

Behavioural data analysis – Hits and FA rates were calculated for each participant, as were three additional scores (see details below), based on the 120 trials in the testing phase.

Eye movement analysis – learning phase – the last trial performed by all participants and the first trial performed by two participants were removed due to technical recording problems. Likewise, six trials of one participant were removed because they were not recorded. As a result, the analysis was based upon 59 trials performed by 55 participants, 58 trials of two participants and 53 trials of one participant, resulting in a total of 3414 analysed trials in this phase.

Testing phase – the analysis was based upon 120 trials performed by 58 participants, minus 12 trials that were removed because no fixations were recorded. As a result, 6948 trials were analysed in this phase.

Results

Behavioural measures – (1) Hit rate, (2) False Alarms (FA) rate, (3) CH score = *p* (Hit) −*p* (FA), (4) Correction for guessing (CFG) [*p* (Hit) − *p* (FA)] / [1−*p* (FA)], 5) (Yonelinas, Regehr, & Jacoby, 1995), Index of signal detection (ISDT) = 1− {4**p* (Hit)* [1 − *p* (FA)]- 2*[*p* (Hit)− *p* (FA)]*[1 + *p* (Hit)− *p* (FA)]}/ {4**p* (Hit)* [1− *p* (FA)]− 2*[*p* (Hit)− p (FA)]*[1 + *p* (Hit)− *p* (FA)]} (Huibregtse, Admiraal, & Meara, 2002).

Eye movements measures – we defined the area of interest as the face itself– 15.1% of the screen on which we calculated these measurements: (1) Dwelling time (ms) – time spent looking at the face (sum of duration of all fixation and saccades). Note that this measure is included only in the testing phase since during the learning phase the face appeared on the screen for a fixed time of three seconds; (2) Fixation rate (number of fixations per second) – Mean number of fixations divided by the duration of the face presentation on the screen; (3) Fixation duration mean (ms) – overall fixation duration divided by number of fixations); (4) First fixation duration mean (ms); (5) Latency to first fixation mean (ms); (6) Total fixation duration (ms) – a measure used to ensure that the time spent looking at subsequent hit trials was identical to time spent looking at miss trials during the learning phase.

Behavioural results – a one-factorial design with group as a between-factor was performed. As can be seen in Table 1, Hit rates did not significantly differ between the Immediate and Delay groups. However, significant differences

Table 1. A one-factorial design (rates of Hits and FAs and scores of CH, ISDT and CFG), with group as a between factor.

Dependent variable	Group	Mean (SD)	*F*
Hit	Immediate	60 (2)	$F(1, 56) = 0.33, p = .57$
	Delay	62 (2)	
FA	Immediate	18 (2)	$F(1, 56) = 4.20, p < .05$
	Delay	24 (2)	
CH	Immediate	41 (2)	$F(1, 56) = 1.55, p = .22$
	Delay	37 (2)	
ISDT	Immediate	44 (2)	$F(1, 56) = 2.87, p < .10$
	Delay	39 (2)	
CFG	Immediate	51 (2)	$F(1, 56) = 0.23, p = .63$
	Delay	49 (3)	

between the groups were found in their FA rates. There was a tendency towards significant differences between the groups on their ISDT scores as well.

Eye movement results

Learning phase

A 2 × 2 mixed-design ANOVA with repeated measures was performed in order to compare the two correctness types (Hit vs. Miss) (within-participants factor) for the two groups (Immediate vs. Delay) (between-participants factor).

Of the total 3414 trials in the learning phase, there were 2074 hit trials (60.75%) and 1340 miss trials (39.25%).

Correctness main effect – we classified the correctness type of the stimuli that appeared in the learning phase as Hit or Miss, based on the participants' results in the testing phase.

Fixation rate – correctness main effect did not reach significance, $F(1, 56) = 3.72$, $p = .06$. There was no significant difference in fixation rate between a Hit trial and a Miss trial. Group main effect did not reach significance, $F(1, 56) = 0.65$, $p = .42$. There was no significant difference between the Immediate group and the Delay group in fixation rate. As can be seen in Figure 1, Correctness by Group interaction reached significance, $F(1, 56) = 4.09$, $p < 0.05$, indicating different eye movement patterns between the two groups in the two types of correctness. *Post hoc* analysis revealed a highly significant difference between Hit and Miss trials in the Immediate group, $F(1, 56) = 8.08$, $p \leq .01$, and no difference between Hit and Miss trials in the Delay group, $F(1, 56) = 0.004$, $p = .95$.

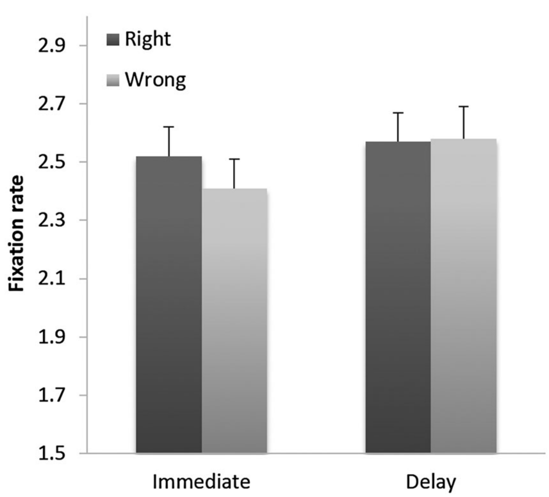

Figure 1. Learning phase – Mean and *SD* of fixation rate per face of the two groups in Accurate/Error trials.

Fixation duration – correctness main effect reached significance, $F(1, 56) = 5.73$, $p < .05$. The fixation duration in a Hit trial ($M = 370.07$, $SE = 18.14$) was shorter than in a Miss trial ($M = 399.6$, $SE = 25.97$). Group main effect did not reach significance, $F(1, 56) = 0.01$, $p = .94$, indicating there was no difference in fixation duration between the Immediate group ($M = 386.56$, $SE = 29.92$) and the Delay group ($M = 383.11$, $SE = 30.97$). Correctness by Group interaction did not reach significance, $F(1, 56) = 0.27$, $p = .61$.

First fixation duration – correctness main effect did not reach significance, $F(1, 56) = 0.57$, $p = .45$, indicating there was no difference between a Hit trial ($M = 403.11$, $SE = 26.16$) and a Miss trial ($M = 388.98$, $SE = 29.73$) in the first fixation duration. Group main effect did not reach significance, $F(1, 56) = 0.12$, $p = .73$, indicating there was no difference between the Immediate group ($M = 405.16$, $SE = 36.67$) and the Delay group ($M = 386.91$, $SE = 37.96$) in the first fixation duration. Correctness by Group interaction did not reach significance either, $F(1, 56) = 0.12$, $p = .73$.

Latency to first fixation – correctness main effect did not reach significance, $F(1, 56) = 0.26$, $p = .61$. There was no significant difference in latency to first fixation between a Hit trial ($M = 45.22$, $SE = 14.87$) and a Miss trial ($M = 48.39$, $SE = 13.89$). Group main effect did not reach significance, $F(1, 56) = 1.09$, $p = .3$ indicating there was no significant difference between the Immediate group ($M = 61.48$, $SE = 19.51$) and the Delay group ($M = 32.14$, $SE = 20.2$) in latency to first fixation. Correctness by Group interaction did not reach significance, $F(1, 56) = 1.53$, $p = .22$.

Total fixation duration – correctness main effect did not reach significance, $F(1, 56) = 0.25$, $p = .62$, indicating there was no difference between a Hit trial ($M = 3205.95$, $SE = 53.37$) and a Miss trial ($M = 3167.74$, $SE = 54.45$).

Testing phase

A 2 × 2×2 mixed-design ANOVA with repeated measures was performed in order to compare the two different correctness types (Accurate – Hits & CR vs. Errors – FA & Miss) (within-participants factor) and the two answer types (Yes – Hits & FA vs. No – CR & Miss) for the two groups (Immediate vs. Delay) (between-participants factor). Of the total 6948 trials in the testing phase, there were 2102 hit trials (30.25%), 2738 CR trials (39.41%), 738 FA trials (10.62%) and 1370 miss trials (19.72%). Average response time (with *SE*(was 2481 (115) ms in hit trials, 2726 (135) ms in CR trials, 2755 (153) ms in FA trials and 2782 (105) ms in miss trials.

Dwelling time – The Answer main effect did not reach significance, $F(1, 56) = 0.77$, $p = .39$, indicating that dwelling time did not differ between a Yes response and a No response. Correctness main effect reached significance,

$F(1, 56) = 36.76$, $p < .001$. Dwelling time in an Accurate trial was shorter than in an Error trial. Group main effect did not reach significance, $F(1, 56) = 0.22$, $p = .64$. There was no difference between the Immediate group and the Delay group in dwelling time. Answer by Group interaction did not reach significance, $F(1, 56) = 0.21$, $p = .65$. Correctness by Group interaction did not reach significance, $F(1, 56) = 0.64$, $p = .43$. As can be seen in Figure 2, Answer by Correctness interaction reached significance, $F(1, 56) = 10.89$, $p < .001$. *Post hoc* analysis revealed no significant difference between Hit and FA rates (Yes answers), $F(1, 56) = 0.20$, $p = .16$, and a significant difference between CR and Miss rates (No answers), $F(1, 56) = 33.7$, $p < .001$. In other words, dwelling time was significantly shorter for CR compared to Miss trials. Answer by Correctness by Group did not reach significance, $F(1, 56) = 1.66$, $p = .2$.

Fixation rate – answer main effect did not reach significance, $F(1, 56) = 0.05$, $p = .83$. The fixation rate did not significantly differ between a Yes response ($M = 2.62$, $SE = 0.06$) and a No response ($M = 2.62$, $SE = 0.05$). As can be seen in Figure 3, correctness main effect reached significance, $F(1, 56) = 4.9$, $p < .05$. There was significant difference in fixation rate between an Accurate trial ($M = 2.6$, $SE = 0.05$) and an Error trial ($M = 2.64$, $SE = 0.06$). Group main effect did not reach significance, $F(1, 56) = 0.78$, $p = .38$, indicating that there was no difference between the Immediate group ($M = 2.57$, $SE = 0.08$) and the Delay group ($M = 2.67$, $SE = 0.08$) in fixation rate. Answer by Group interaction did not reach significance, $F(1, 56) = 0.03$, $p = .87$. Correctness by Group interaction did not reach significance, $F(1, 56) = 0.001$, $p = .97$. Answer by Correctness interaction was marginally significant, $F(1, 56) = 2.98$, $p = .09$, as was the Answer by Correctness by Group interaction, $F(1, 56) = 3.41$, $p = .07$. These results indicate that in the Delay group, as opposed to the Immediate group, the fixation rate between correct and incorrect responses was lower for No answers than for Yes answers.

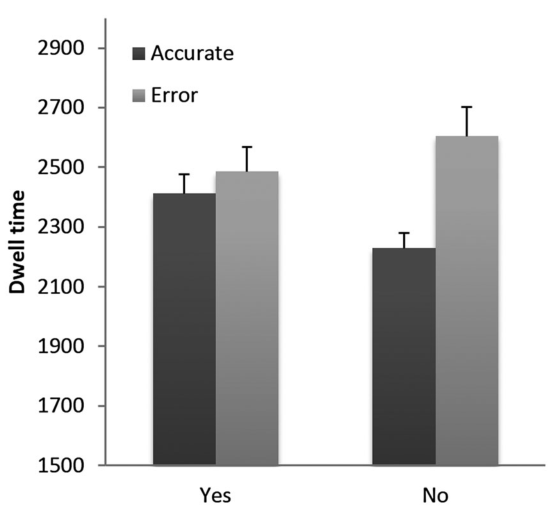

Figure 2. Testing phase – Mean and *SD* of dwelling time per face for the two groups combined in Accurate/Error trials as a function of a yes or no answer.

Fixation duration – Answer main effect did not reach significance, $F(1, 56) = 0.41$, $p = .52$. There was no significant difference in fixation duration between a Yes answer and a No answer. Correctness main effect did not reach significance, $F(1, 56) = 1.39$, $p = .24$ There was no significant difference in fixation duration between an Accurate trial and an Error trial. Group main effect did not reach significance, $F(1, 56) = 0.1$, $p = .75$, indicating there was no significant difference between the Immediate group and the Delay group in fixation duration. Answer by Group interaction did not reach significance, $F(1, 56) = 0.51$, $p = .48$. Correctness by Group interaction did not reach significance, $F(1, 56) = 0.15$, $p = .7$. As can be seen in Figure 4(a), Answer by Correctness interaction reached significance, $F(1, 56) = 4.93$, $p < .05$. *Post hoc* analysis revealed no significant difference between Accurate and Error trials in a Yes answer $F(1, 56) = 0.63$, $p = .43$ and a significant difference between an Accurate and an Error trial in a No answer $F(1, 56) = 10.18$, $p < .01$. Answer by Correctness by Group reached significance, $F(1, 56) = 4.44$, $p < .05$. *Post hoc* analysis did not reveal significant Answer by Correctness interaction in the Immediate group (see Figure 4b). Yes answers yielded no differences between an Accurate trial and an Error trial, $F(1, 56) = 0.07$, $p = .8$. No answers yielded no difference between an Accurate trial and an Error trial, $F(1, 56) = 0.31$, $p = .58$. In the Delay group there was an Answer by Correctness significant interaction (see Figure 4c). In a Yes answer there was no significant difference between an Accurate trial and an Error trial, $F(1, 56) = 1.83$, $p = .18$. In a No answer there was a significant difference between an Accurate trial and an Error trial, $F(1, 56) = 15.22$, $p < 0.001$.

First fixation duration – answer main effect did not reach significance, $F(1, 56) = 0.11$, $p = .75$. There was no significant difference between a Yes response ($M = 321.77$, $SE = 21.99$) and a No response ($M = 293.11$, $SE = 22.76$) in the first fixation duration. Correctness main effect did not reach significance, $F(1, 56) = 0.56$, $p = .46$. There was no significant difference between an Accurate trial ($M = 303.16$, $SE = 13.16$) and an Error trial ($M = 311.727$, $SE = 19.82$) in the first fixation duration. Group main effect did not reach significance, $F(1, 56) = 0.82$, $p = .36$. There was no difference between the Immediate group ($M = 321.77$, $SE = 21.99$) and the Delay group ($M = 293.11$, $SE = 22.76$) in the first fixation duration. Answer by Group interaction did not reach significance, $F(1, 56) = 1.92$, $p = .17$. Correctness by Group interaction did not reach significance, $F(1, 56) = 0.04$, $p = .84$. Answer by Correctness interaction did not reach significance, $F(1, 56) = 3.73$, $p = .06$. Answer by Correctness by Group did not reach significance, $F(1, 56) = 0.003$, $p = .96$.

Latency to first fixation – answer main effect did not reach significance, $F(1, 56) = 0.22$, $p = .64$, indicating there was no significant difference between a Yes answer and a No answer in latency to first fixation. Correctness main effect did not reach significance, $F(1, 56) = 0.01$, $p = .93$, indicating there was no significant difference between an

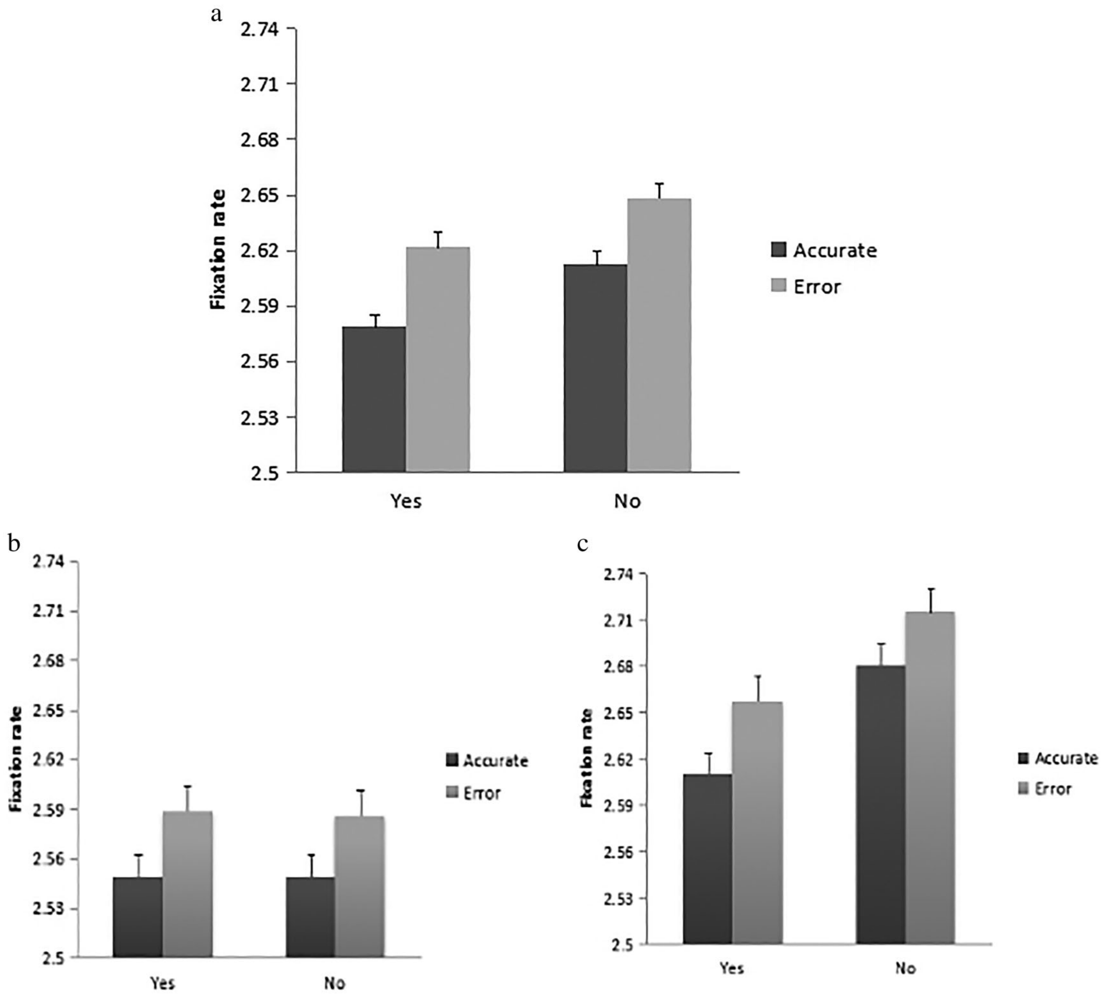

Figure 3. (a) Testing phase – Mean and *SD* of fixation rate per face for the two groups combined in Accurate/Error trials as a function of their yes or no answer. (b) Testing phase – Mean and *SD* of fixation rate per face for the two types of answers in Accurate/Error trials of the Immediate group. (c) Testing phase – Mean and *SD* of fixation rate per face for the two types of answers in Accurate/Error trials of the Delay group.

Accurate trial and an Error trial in latency to first fixation. Group main effect did not reach significance, $F(1, 56) = 0.61$, $p = .44$, indicating there was no significant difference between the Immediate group and the Delay group in latency to first fixation. Answer by Group interaction did not reach significance, $F(1, 56) = 0.15$, $p = .7$. As can be seen in Figure 5, Correctness by Group interaction reached significance, $F(1, 56) = 4.26$, $p < .05$, although *post hoc* analysis did not reveal a significant difference between an Accurate trial and an Error trial in the Immediate group, $F(1, 56) = 2.03$, $p = .16$ or in the Delay group, $F(1, 56) = 2.23$, $p = .14$. Answer by Correctness interaction did not reach significance, $F(1, 56) = 0.02$, $p = .9$. Answer by Correctness by Group interaction did not reach significance, $F(1, 56) = 0.4$, $p = .53$.

Discussion

This study examined eye movements in Hit and Miss results during the learning phase, and Hit, Miss, CR and FA trials during the testing phase in a yes/no recognition paradigm of pictures of human faces. Current literature shows that Hit, Miss, CR and FA trials are related to different cognitive and brain processes (Mecklinger, 2000; Osipova et al., 2006). In addition, human eye movements can predict recognition processes (Hannula, Althoff, et al., 2010; Hannula, Ranganath, et al., 2010) and can affect future recognition (Heisz et al., 2013). The novelty of this study is that, to the best of our knowledge, it is the first to integrate these aspects and to test eye movements in all four conditions, thus expanding knowledge on the role of eye movements in recognition tests. It is also the first to examine the effect of time delay on eye movements in all the four conditions.

The two groups of participants differed in the intervals given between the learning and testing phases (i.e. immediate and 15-minute delay). Consistent with our predictions and prior findings (Chance & Goldstein, 1987), the Delay group had a higher FA rate, and there was no difference in Hit rates between the two groups. The CH scores

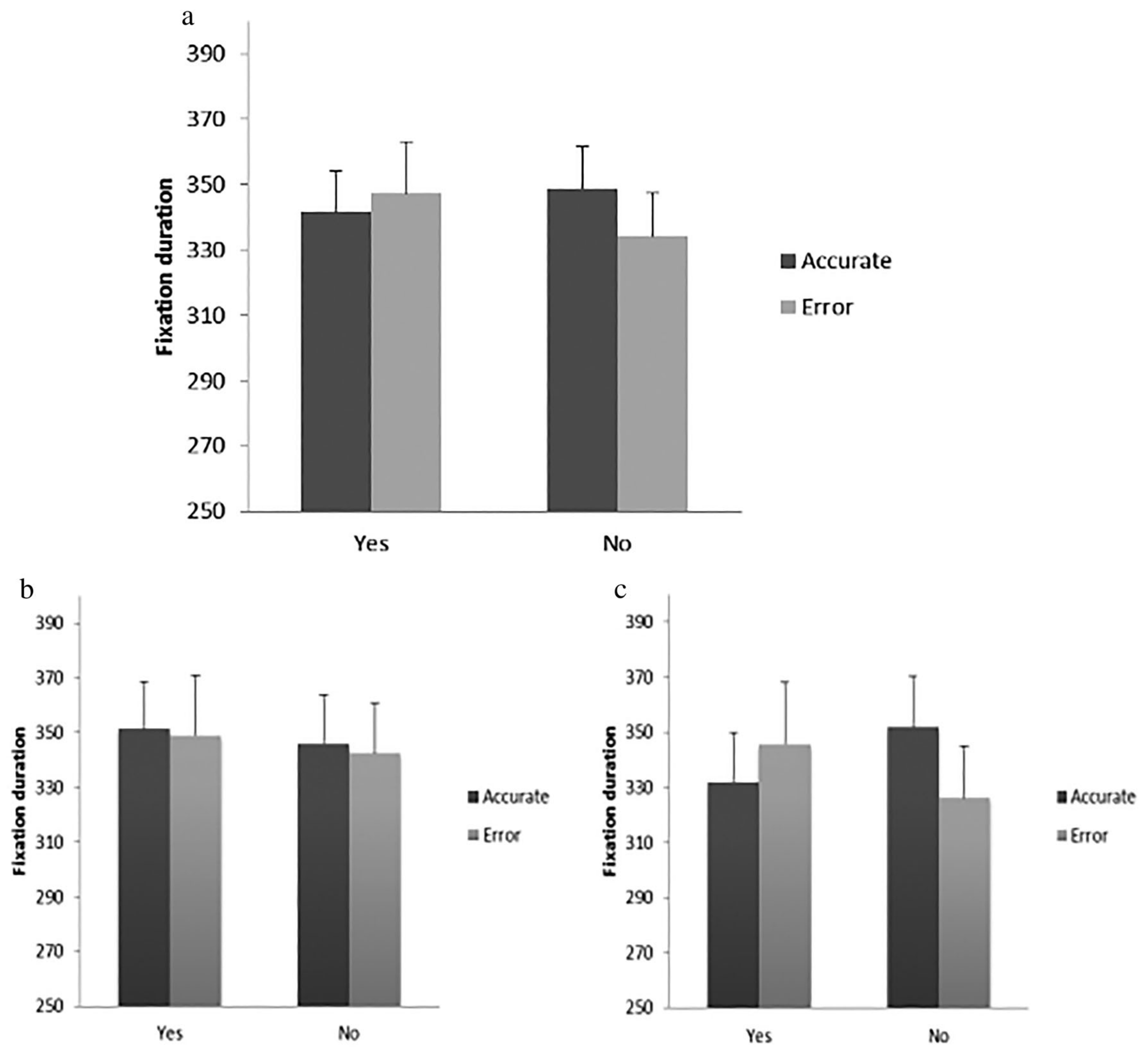

Figure 4. (a) Testing phase – Mean and *SD* of fixation duration per face of the two types of answers in Accurate/Error trials. (b) Testing phase – Mean and *SD* of fixation duration per face of the two types of answers in Accurate/Error trials of the Immediate group. (c) Testing phase – Mean and *SD* of fixation rate per face of the two groups in Accurate/Error trials of the Delay group.

(Roediger & McDermott, 1994) of the two groups were not significantly different either. We applied two additional indices for calculating recognition accuracy. The first index is Correction for Guessing (CFG) (Yonelinas et al., 1995) and the second is Index of Signal Detection (ISDT) (Huibregtse et al., 2002). The Immediate group tended to be more accurate than the Delay group, as revealed by the ISDT index.

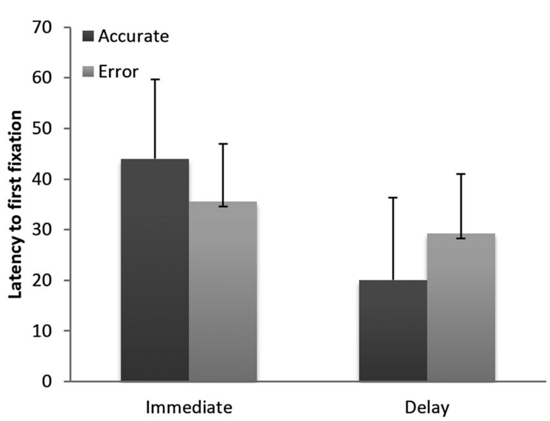

Figure 5. Testing phase – Mean and *SD* of latency to first fixation per face of the two groups in Accurate/Error trials.

As predicted, significant Correctness main effects were found in the learning and testing phases. Specifically, Hit responses had shorter mean fixation durations in the learning phase than Miss responses (the meaning of these results cannot be explained by the difference in time spent looking at the screen in Hit vs. Miss trials during the learning phase since no such difference was found, as measured by total fixation duration). Also, Accurate trials (Hit and CR) had shorter dwelling times and fewer fixations than Error trials (Miss and FA) in the testing phase.

As predicted, significant Answer by Correctness interactions were shown. The difference in dwelling time between an Accurate trial and an Error trial was significant only for a No answer (Figures 2–4), with a tendency towards significance between the two trials for fixation duration. Also, Accurate trials had shorter mean fixation durations than Error trials in a Yes answer and longer mean fixation durations than an Error trial in a No answer (Figure 4a).

Contrary to our prediction, the Immediate group did not display a stronger "repetition effect" than the Delay group (fewer fixations and longer first and overall fixation duration mean). However, this lack of effect is perhaps due to low accuracy rate – 60 percent for the Immediate group and 62 percent for the Delay group. In other studies that showed the repetition effect – either the analysis examined correct trials only (since accuracy level is almost 100%) (Chanon & Hopfinger, 2008), or accuracy level is much higher than in our study (Ryan et al., 2007). Nonetheless, the repetition effect was indeed manifested in the fact that during the testing phase, Accurate trials received less fixations and shorter dwelling time than Error trials.

The pattern of results of our study, in which no Answer main effect reached significance and no Answer by Group interaction reached significance, coincides with data in recent literature that suggests that eye movements are not indicative of the participant's behavioural decisions (Hannula et al., 2012). This is especially impressive when taking into account that there were Answer by Correctness significant interactions (Figures 2 –4) and an Answer by Correctness by Group significant interaction (Figure 4(b, c)).

In addition, there were Correctness by Group interactions at the learning and testing phases. (1) During the learning phase, the difference in fixation rates between the Hit and Miss trials was significant only in the Immediate group. Specifically, a Hit trial received more fixations than a Miss trial in the Immediate group (Figure 1). (2) During the testing phase, the latency to first fixation was longer in an Accurate trial in the Immediate group, but shorter in the Delay group when compared with Error trials (Figure 5).

Regarding the relationship between each of the measures and the attention component to which they are related, as mentioned in the introduction (Chanon & Hopfinger, 2008; Lai et al., 2013), faces that were remembered correctly were those that – (a) received more attention during the learning phase (although actual significance was only shown for the Immediate group, while only a trend that almost reached significance was shown for the Delay group); (b) received less attention (shorter fixation) in each fixation during the learning phase as compared to Error trials. The latter effect (shorter fixation at the learning phase) is not due to the fixed window of face appearance duration at the learning phase (one might claim that the higher fixation mean may have been at the expense of the length of duration of each fixation), since the fixation rate mean was calculated as the number of fixations per second; (c) received less attention (as reflected by their significantly lower fixation rate mean) during the testing phase. These conclusions could not have been obtained had we used solely behavioural measures, without eye movement recordings.

In summary, our study revealed three major findings. First, although eye movements diverge from the behavioural decision, (yes versus no), as reflected by the lack of significant Answer main effects and Answer by Group interactions, the current results indicate that eye movements do align well with memory accuracy and with memory accuracy within each type of decision. This is reflected by significant Correctness main effects during both phases and by the significant Answer by Correctness interaction during the testing phase. These results also coincide with current literature by showing a lack of association between explicit response and eye movements (Hannula et al., 2012), though there is an association between eye movements and decision making (Orquin & Loose, 2013). Consistent with this literature, the participant's answer in this current study was associated with eye movements but only through interaction with correctness type and not just in itself. The results of the present study highlight the need to use both Correctness type and Answer type as independent variables.

The second major finding is that although there was no Group main effect at the testing phase, the time interval between the learning and testing phases did interact with eye movements, as reflected by Correctness by Group interactions at the learning and testing phases. These results coincide with data in the current literature that suggests that time delay has an effect on the recognition processes (Schacter, 1999). In other words, our study showed that the known effect of time delay on the recognition processes is also reflected by eye movements. Also, these results emphasise the need for caution when interpreting results regarding eye movements without controlling the time interval between learning and testing phases.

The third contribution of the present study is that it expands the findings about relationships between eye movements and recognition processes. It was shown that there is disproportionate viewing time in favour of matching faces, before overt recognition (Hannula and Ranganath 2009). However, our study expanded these findings and examined this relationship using various indicators that are related to indices of eye movements. This is a major qualitative difference since many studies have shown that indicators related specifically to fixation patterns (that were used in this study), cannot be controlled by participants. For example, even if participants tried to conceal their familiarity with the stimuli, their fixation patterns, such as fixation duration and number of fixations, would "reveal" it (Peth, Kim, & Gamer, 2013; Schwedes & Wentura, 2016; Schwedes & Wentura, 2012).

In conclusion, this study clearly demonstrated the added value of the information generated from participants' eye movements at the learning phase as well as at the testing phase, over behavioural information.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abdi, H. (2007). Signal detection theory. In N.J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 886–889). Thousand Oaks, CA: Sage.

Barton, J. J., Radcliffe, N., Cherkasova, M. V., & Edelman, J. A. (2007). Scan patterns at the processing of facial identity in prosopagnosia. *Experimental Brain Research*, *181*, 199–211. doi:10.1007/s00221-007-0923-2

Chance, J. E. & Goldstein, A. G. (1987). Retention interval and face recognition: Response latency measures. *Bulletin of the Psychonomic Society*, *25*(6), 415–418. doi:10.3758/BF03334728

Chanon, V. & Hopfinger, J. B. (2008). Memory's grip on attention: The influence of item memory on the allocation of attention. *Visual Cognition*, *16*, 325–340. doi:10.1080/13506280701459026

Colbert, J. M., & McBride, D. M. (2007). Comparing decay rates for accurate and false memories in the DRM paradigm. *Memory & Cognition*, *35*(7), 1600–1609. doi:10.3758/BF03193494

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: Using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*, 1–16. doi:10.3389/fnhum.2010.00166

Hannula, D. E., Baym, C. L., Warren, D. E., & Cohen, N. J. (2012). The eyes know eye movements as a veridical index of memory. *Psychological Science*, *23*(3), 278–287. doi:10.1177/0956797611429799

Hannula, D. E. & Ranganath, C. (2009). The eyes have it: Hippocampal activity predicts expression of memory in eye movements. *Neuron*, *63*(5), 592–599. doi:10.1016/j.neuron.2009.08.025

Hannula, D. E., Ranganath, C., Ramsay, I. S., Solomon, M., Yoon, J., Niendam, T. A., & Ragland, J. D. (2010). Use of eye movement monitoring to examine item and relational memory in schizophrenia. *Biological Psychiatry*, *68*(7), 610–616. doi:10.1016/j.biopsych.2010.06.001

Harrington, M. (2006). The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing*, *23*(1), 73–98. doi:10.1191/0265532206lt321oa

Heisz, J. J., Pottruff, M. M., & Shore, D. I. (2013). Females scan more than males A potential mechanism for Sex differences in recognition memory. *Psychological Science*, *24*, 1157–1163. doi:10.1177/0956797612468281

Huibregtse, I., Admiraal, W., & Meara, P. (2002). Scores on a yes–no vocabulary test: Correction for guessing and response style. *Language Testing*, *19*(3), 227–245. doi:10.1191/0265532202lt229oa

Kafkas, A. & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *The Quarterly Journal of Experimental Psychology*, *64*, 1971–1989. doi:10.1080/17470218.2011.588335

Koen, J. D., Aly, M., Wang, W.-C., & Yonelinas, A. P. (2013). Examining the causes of memory strength variability: Recollection, attention failure, or encoding variability? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(6), 1726–1741. doi:10.1037/a0033671

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., … Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115. doi:10.1016/j.edurev.2013.10.001

Laughery, K. R., Fessler, P. K., Lenorovitz, D. R., & Yoblick, D. A. (1974). Time delay and similarity effects in facial recognition. *Journal of Applied Psychology*, *59*, 490–496. doi:10.1037/h0037193

MacLin, O. H., MacLin, M. K., & Malpass, R. S. (2001). Race, arousal, attention, exposure and delay: An examination of factors moderating face recognition. *Psychology, Public Policy, and Law*, *7*(1), 134–152. doi:10.1037/1076-8971.7.1.134

Macmillan, N. A. (2002). Signal detection theory. In H. Pashler (Ed.), *Stevens' handbook of experimental psychology* (pp. 43–91). New York, NY: John Wiley & Sons, Inc.

Mecklinger, A. (2000). Interfacing mind and brain: A neurocognitive model of recognition memory. *Psychophysiology*, *37*, 565–582. doi:10.1111/1469-8986.3750565

Messer, K., Matas, J., Kittler, J., Luettin, J., & Maitre, G. (1999). *XM2VTSbd: The extended M2VTS database, proceedings 2nd conference on audio and video-base biometric personal verification (AVBPA99)*. Springer, New York, NY.

Mitchell, D. B., Brown, A. S., & Murphy, D. R. (1990). Dissociations between procedural and episodic memory: Effects of time and aging. *Psychology and Aging*, *5*(2), 264–276. doi:10.1037/0882-7974.5.2.264

Orquin, J. L. & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206. doi:10.1016/j.actpsy.2013.06.003

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *The Journal of Neuroscience*, *26*(28), 7523–7531. doi:10.1523/JNEUROSCI.1948-06.2006

Peth, J., Kim, J. S., & Gamer, M. (2013). Fixations and eye-blinks allow for detecting concealed crime related memories. *International Journal of Psychophysiology*, *88*(1), 96–103. doi:10.1016/j.ijpsycho.2013.03.003

Roediger, H. L. & McDermott, K. B. (1994). The problem of differing false-alarm rates for the process dissociation procedure: Comment on Verfaellie and Treadwell (1993). *Neuropsychology*, *8*, 284–288. doi:10.1037/0894-4105.8.2.284

Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, *11*, 454–461. doi:10.1111/1467-9280.00288

Ryan, J. D., Hannula, D. E., & Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory (Hove, England)*, *15*, 508–525. doi:10.1080/09658210701391022

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*(3), 182–203. doi:10.1037/0003-066X.54.3.182

Schwedes, C. & Wentura, D. (2012). The revealing glance: Eye gaze behavior to concealed information. *Memory & Cognition*, *40*(4), 642–651. doi:10.3758/s13421-011-0173-1

Schwedes, C., & Wentura, D. (2016). Through the eyes to memory: Fixation durations as an early indirect index of concealed knowledge. *Memory & Cognition*, *44*(8), 1244–1258. doi:10.3758/s13421-016-0630-y

Sharot, T., Davidson, M. L., Carson, M. M., & Phelps, E. A. (2008). Eye movements predict recollective experience. *PLoS One*, *3*, e2884. doi:10.1371/journal.pone.0002884

Shepherd, J. W., Gibling, F., & Ellis, H. D. (1991). The effects of distinctiveness, presentation time and delay on face recognition. *European Journal of Cognitive Psychology*, *3*(1), 137–145. doi:10.1080/09541449108406223

Slotnick, S. D. & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, *7*, 664–672. doi:10.1038/nn1252

Smith, C. N. & Squire, L. R. (2008). Experience-dependent eye movements reflect hippocampus-dependent (aware) memory. *The Journal of Neuroscience*, *28*, 12825–12833. doi:10.1523/JNEUROSCI.4542-08.2008

Snodgrass, J. G. & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50. doi:10.1037/0096-3445.117.1.34

Stacey, P. C., Walker, S., & Underwood, J. D. (2005). Face processing and familiarity: Evidence from eye-movement data. *British Journal of Psychology*, *96*(4), 407–422. doi:10.1348/000712605X47422

Stanislaw, H. & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. doi:10.3758/BF03207704

Tanaka, J. W. & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46*(2), 225–245. doi:10.1080/14640749308401045

Thapar, A. & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition*, *29*(3), 424–432. doi:10.3758/BF03196393

Tulving, E. & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373. doi:10.1037/h0020071

Yonelinas, A. P., Regehr, G., & Jacoby, L. L. (1995). Incorporating response bias in a dual-process theory of memory. *Journal of Memory and Language*, *34*(6), 821–835. doi:10.1006/jmla.1995.1036