COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 41 (2017) 1533–1554

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12437

Spatial But Not Oculomotor Information Biases Perceptual Memory: Evidence From Face Perception and Cognitive Modeling

Andrea L. Wantz,[a,b] Janek S. Lobmaier,[a,b] Fred W. Mast,[a,b] Walter Senn[b,c]

[a]*Department of Psychology, University of Bern*
[b]*Center for Cognition, Learning and Memory, University of Bern*
[c]*Department of Physiology, University of Bern*

Received 26 March 2015; received in revised form 2 February 2016; accepted 30 June 2016

Abstract

Recent research put forward the hypothesis that eye movements are integrated in memory representations and are reactivated when later recalled. However, “looking back to nothing” during recall might be a consequence of spatial memory retrieval. Here, we aimed at distinguishing between the effect of spatial and oculomotor information on perceptual memory. Participants’ task was to judge whether a morph looked rather like the first or second previously presented face. Crucially, faces and morphs were presented in a way that the morph reactivated oculomotor and/or spatial information associated with one of the previously encoded faces. Perceptual face memory was largely influenced by these manipulations. We considered a simple computational model with an excellent match (4.3% error) that expresses these biases as a linear combination of recency, saccade, and location. Surprisingly, saccades did not play a role. The results suggest that spatial and temporal rather than oculomotor information biases perceptual face memory.

Keywords: Eye movements; Memory; Perceptual memory; Cognitive model

1. Introduction

Most people will have experienced a situation in which they misplaced an object, for example, keys or glasses, in absent-mindedness. One common strategy to find lost objects is to turn around and retrace the path on which the object might have gotten lost. In memory tasks, by applying the “method of loci,” participants mentally place to be

---

Correspondence should be sent to Andrea Laura Wantz, Department of Psychology and Center for Cognition, Learning and Memory, University of Bern, Fabrikstrasse 8, CH-3012 Bern, Switzerland. E-mail: andrea.wantz@gmail.com

memorized material at locations on a familiar path, for example, the way home from work. During recall, the path is mentally rewalked to facilitate memory recall. Similarly, it has been demonstrated that the eyes "walk back" to the encoding location when we recall or mentally visualize a previously seen object (Johansson & Johansson, 2014; Laeng & Teodorescu, 2002; Martarelli & Mast, 2013; Spivey & Geng, 2001). Here, we ask whether looking back to a location where information has been encoded biases perceptual face memory.

Some studies have demonstrated that looking back to the encoding location during memory retrieval is functional for recall. For example, memory recall is impaired during fixation compared to free viewing (Johansson, Holsanova, Dewhurst, & Holmqvist, 2012; Laeng, Bloem, D'Ascenzo, & Tommasi, 2014; Laeng & Teodorescu, 2002) and when the eyes are guided to a position incongruent with the encoding location compared to congruent locations (Johansson & Johansson, 2014; but see Martarelli & Mast, 2013). These studies support the hypothesis that eye movements are stored in integrated memory representations and are co-activated when memories are retrieved (Laeng & Teodorescu, 2002).

Interestingly, during recall or mental imagery, the eyes do not only look back to the encoding location but sometimes reenact the entire "scan-path" produced during stimulus viewing (Brandt & Stark, 1997; Spivey & Geng, 2001). Similarly, scanpaths during encoding and recognition (Foulsham & Underwood, 2008) or image re-viewing (Harding & Bloj, 2010; Humphrey & Underwood, 2008; Underwood, Foulsham, & Humphrey, 2009) often closely match. The "scanpath theory" proposed by Noton and Stark (1971) predicts that during pattern recognition, the sequence of fixations is compared to the stored sequence enacted during encoding. When the paths match, the pattern is recognized. This theory rests upon early ideas of Hebb (1968), namely, that a mental image is not generated at once but in parts, fixation after fixation.

Recently, Foulsham and Kingstone (2013) conducted a study to experimentally test the scanpath theory. Participants encoded scenes and later performed a recognition task. Although participants looked back to fixated positions during recognition, they did not do so in the same temporal order (Experiment 1). Moreover, recognition accuracy did not differ between trials on which scanpaths were reenacted in the same compared to different temporal order as during encoding (Experiment 5), hence emphasizing the role of spatial information during recognition memory.

Such spatial information might explain why previous studies found that during recall, people perform eye movements back to the encoding location. For instance, if an elephant was encoded in the lower right quadrant of the screen, participants will be more likely to look back into this area when later recalling this animal on a blank screen (Johansson & Johansson, 2014). To date, it remains unclear whether this looking back is based on the reactivation of oculomotor information that is integrated into memory representations or, alternatively, whether these eye movements are a consequence of spatial memory recall.

In this study we aimed at disentangling effects of oculomotor and spatial information on perceptual memory using a novel task. The basic principle of our task lies in the assumption that a memory trace of previously encoded information influences the perception of an ambiguous stimulus. On each trial, participants were presented with two different faces at two different locations on the screen. Then, a morph consisting of the exact average of the two faces appeared and participants were to judge whether it looked more like the first or second face. Crucially, the morph was the exact average of the two faces so that there were no correct or wrong responses. We expected that participants' judgments would be biased by the experimental manipulations, that is, by the reactivation of oculomotor and/or spatial information associated with one of the two faces during encoding. Oculomotor information was reactivated by presenting fixation crosses and morphs in a way that participants enacted a similar saccade as during encoding of specific face while spatial information was reactivated by presenting the morph at the very location where a specific face has been encoded. We propose a computational model expressing these response biases as a function of individual contributions of oculomotor information (saccades), spatial information (location), and recency.

2. Methods

2.1. Participants

A total of 153 participants were tested. Each participant was tested under only one of the four conditions. Prior to the experiment, all participants gave written informed consent to participate. Participants were treated in accordance to the protocol approved by the Faculty of Human Science of the University of Bern and conformed to the "Ethical Principles of Psychologists and Code of Conduct" of the American Psychological Association (2002).

2.2. Material

A total of 96 gray-scale face images were derived from the FERET Database (Phillips, Moon, Rizvi, & Rauss, 2000; Phillips, Wechsler, Huang, & Rauss, 1998). The 96 images were arranged in pairs of the same gender and approximately the same age. We then created the exact average of each pair by morphing the faces using PsychoMorph computer graphics software (Burt & Perrett, 1995; Tiddeman, Burt, & Perrett, 2001). All images were aligned by the eyes and cropped ovally using Photoshop. The final image subtended a visual angle of approximately $7 \times 8°$ (image size of 6×7 cm at a viewing distance of 50 cm). To suppress afterimages of the faces, a mask (see Fig. 1) was created for each face from the image itself using a Matlab script. Fixation crosses were presented in black 36pt Courier font.

2.3. Procedure

All experimental conditions were programmed in E-Prime (Psychology Software Tools Inc., Pittsburgh, PA, USA; http://www.pstnet.com/prime) and were run on personal

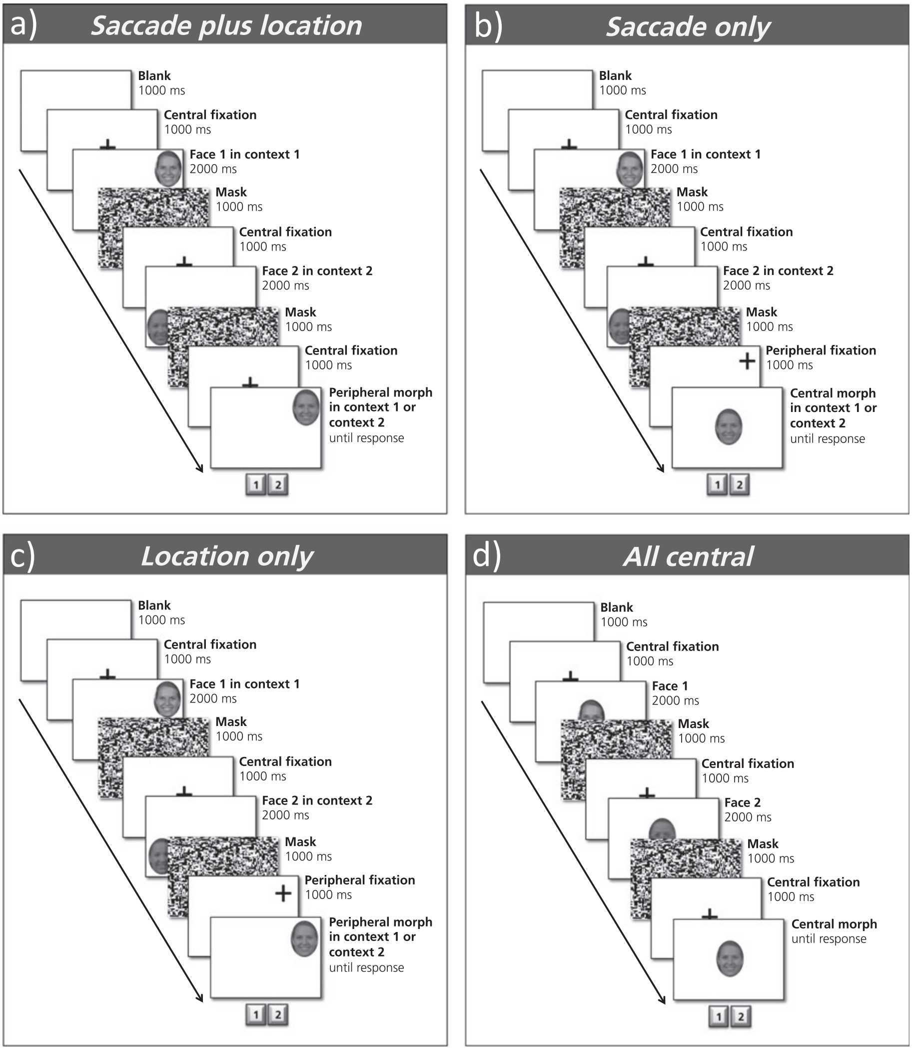

Fig. 1. Experimental procedures of the four conditions. On each trial, participants were presented with a first face in Context 1 (specific saccades and/or locations), a second face in Context 2, and the average of them (morph) with (a) saccades and location, (b) only the saccade, or (c) only the location of either Context 1 or Context 2. Finally (d), all face images were presented centrally without triggering saccades. In all four conditions, participants were asked to judge whether the morph looks more like the first or second face.

laptops. Participants were assigned to one of four conditions and were told that the study is about face perception. The instruction was the same for all conditions. Participants were informed that two faces would be presented after each other, which they were to memorize. Then, a third face would be presented and their task would be to judge whether it looks more like the first or second face. Responses were given by pressing key “1” on the keyboard for first face responses and key “2” for second face responses while response time was not emphasized. Importantly, participants were encouraged to always fixate the fixation cross between the faces as the fixation cross determined saccades. As described in Appendix A, it is likely that participants followed this instruction.

Each of the four conditions comprised 48 trials. One trial consisted of nine slides: a blank screen for 1,000 ms, a central fixation cross for 1,000 ms, the presentation of the first face for 2,000 ms, a mask for 1,000 ms, a central fixation cross for 1,000 ms, the presentation of the second face for 2,000 ms, a mask for 1,000 ms, a fixation cross for 1,000 ms, and finally the presentation of the morph until a response was given. Throughout the experiment, the background was white. The procedures of all four conditions are illustrated in Fig. 1. The same pairs were used in all four conditions but appeared in a random order. For each pair, the face that appeared first on each trial was randomly chosen.

2.3.1. Saccade plus location condition

Forty-four participants (28 women, $M_{\text{age}} = 32.2$ years, $SD = 15.87$, range = 19–70) were tested in this condition. In the *saccade plus location* condition, the morph was presented in a way that both the location and the saccade to one of the encoded faces were repeated. Before encoding each of the two faces at a different location on the screen, participants had to fixate a central fixation cross to execute a specific saccade. The morph was then presented at either of the two locations and was also preceded by a central fixation cross. In half of the trials the morphed face appeared in the context of the first face (same saccade and location as the first face), in the other half of the trials it appeared in the context of the second face (same saccade and location as the second face). We expected a higher probability for participants to judge the morph as looking more similar to the face that was encoded with the same saccade and at the same location compared to the other face.

2.3.2. Saccade-only condition

Thirty-two participants were tested in the *saccade-only* condition (15 women, $M_{\text{age}} = 34.25$ years, $SD = 16.25$, range = 18–67). In this condition, we presented the morph so that one of the two saccades executed during encoding was repeated but the morph appeared in a new location. Thus, on each trial the two faces appeared at a peripheral location on the screen preceded by a central fixation cross. The morph then appeared centrally at all times but was preceded by a peripheral fixation cross in a way that a saccade enacted during encoding was repeated. For instance, if the first face was encoded in the upper right corner, participants executed a saccade from the center diagonally toward the upper right corner during encoding. If the morph should then be presented with the same oculomotor vector, a fixation cross was presented in the lower left corner and the morph appeared centrally. In half of the trials, the morph was presented with the saccade enacted during encoding the first face (i.e., in Context 1) without repeating the location. In the remaining trials, only the saccade but not the location associated with the second face was repeated. We expected a higher probability for participants to judge the morph as resembling the face that was encoded with the same saccade.

2.3.3. Location-only condition

Forty-three participants (24 women, $M_{\text{age}} = 29.88$ years, $SD = 10.76$, range = 18–60) were tested in the location-only condition. Here, the morph was presented either at the same location as the first or second face, whereas no saccade was executed immediately beforehand. Thus, like in the previously described conditions, two faces were encoded at different locations on the screen, preceded by a central fixation cross, so that saccades were executed. The morph then appeared either in the location of the first or second face, but crucially, the preceding fixation cross appeared in exactly the same position as the morph did. Thus, participant's eyes were already at the very location where the morph was presented and no saccade was elicited. In half of the trials, the morph was presented in the location of the first face (i.e., Context 1), in the other half the morph was presented in the location of the second face (i.e., Context 2). We expected a higher probability for participants to judge the morph as the face that was encoded in the same location.

2.3.4. Recency only

Finally, 34 participants were tested in this condition (24 women, $M_{\text{age}} = 24.29$, $SD = 5.04$, range = 19–45). To control for biases other than spatial or oculomotor influences, all faces and fixation crosses as well as the morph appeared in the center of the screen in this condition. As saccades and location did not vary between encoded faces, we expected participants' judgments to only depend on recency (a generally higher proportion of second face responses compared to first face responses).

3. Results

As the morphs were comprised of exactly 50% of each previously encoded face, there was no correct response for perceptual judgments. Instead, we were interested in the proportion of biased responses, that is, the probability of judging the morph as resembling Face i when the context was repeating both saccade and location, only saccade, or only location of the face presentation i ($i = 1, 2$). In the *recency-only* condition, responses were not biased by spatial or oculomotor information, but only by recency. The descriptive data are summarized in Table 1A.

3.1. Statistical analysis

To analyze differences in response biases depending on the context (experimental condition), a mixed-effects analysis of variance (ANOVA) was computed with Face (first face and second face) as a within-participants factor and condition (*saccade plus location*, *saccade only*, *location only*, and *recency only*) as a between-participants factor. The results revealed significant main effects of face ($F(1, 149) = 207.287$, $p < .001$, $\eta_p^2 = 0.582$) and of condition ($F(3, 149) = 37.811$, $p < .001$, $\eta_p^2 = 0.432$). The main effect of face indicates more biased responses for Face 2 compared to Face 1. Tukey corrected post hoc comparisons revealed more biased responses in the *saccade plus location* condition

Table 1
Empirical choice probabilities (standard deviations) and choice probabilities for face identity unit *i* given context *c* as predicted by the logistic regression function and the proposed computational model

Experiment	Condition	A) Empirical Proportions (SD)	B) Activation Probability Logistic Regression	C) Activation Probability Neuronal Model
Saccade	$p_1(\mathrm{fld}_1 \mid \mathrm{context}_1)$	0.479 (0.166)	$P(\mathrm{fld}_1 \mid c_1) = \sigma\,(-\,r + s + l)$	$P(\mathrm{fld}_1 \mid c_1) = 0.5 - r + s + l$
plus location	$p_2(\mathrm{fld}_2 \mid \mathrm{context}_2)$	0.757 (0.124)	$P(\mathrm{fld}_2 \mid c_2) = \sigma\,(r + s + l)$	$P(\mathrm{fld}_2 \mid c_2) = 0.5 + r + s + l$
Saccade only	$p_1(\mathrm{fld}_1 \mid \mathrm{context}_1)$	0.284 (0.124)	$P(\mathrm{fld}_1 \mid c_1{}^{-l}) = \sigma\,(-\,r + s)$	$P(\mathrm{fld}_1 \mid c_1{}^{-l}) = (0.5 - r + s - l)/(1\text{-}2l)$
	$p_2(\mathrm{fld}_2 \mid \mathrm{context}_2)$	0.681 (0.141)	$P(\mathrm{fld}_2 \mid c_2{}^{-l}) = \sigma\,(r + s)$	$P(\mathrm{fld}_2 \mid c_2{}^{-l}) = (0.5 + r + s - l)/(1\text{-}2l)$
Location only	$p_1(\mathrm{fld}_1 \mid \mathrm{context}_1)$	0.516 (0.135)	$P(\mathrm{fld}_1 \mid c_1{}^{-\mathrm{s}}) = \sigma\,(-\,r + l)$	$P(\mathrm{fld}_1 \mid c_1{}^{-\mathrm{s}}) = (0.5 - r - s + l)/(1\text{-}2s)$
	$p_2(\mathrm{fld}_2 \mid \mathrm{context}_2)$	0.678 (0.146)	$P(\mathrm{fld}_2 \mid c_2{}^{-\mathrm{s}}) = \sigma\,(r + l)$	$P(\mathrm{fld}_2 \mid c_2{}^{-\mathrm{s}}) = (0.5 + r - s + l)/(1\text{-}2s)$
Recency only	$p_1(\mathrm{fld}_1 \mid \mathrm{context}_1)$	0.382 (0.099)	$P(\mathrm{fld}_1 \mid c_0{}^{+l}) = \sigma\,(-\,r)$	$P(\mathrm{fld}_1 \mid c_0{}^{+l}) = (0.5 - r + l)/(1{+}2l)$
	$p_2(\mathrm{fld}_2 \mid \mathrm{context}_2)$	(1−0.382 = 0.618)	$P(\mathrm{fld}_2 \mid c_0{}^{+l}) = 1 - P(\mathrm{fid}_1 \mid c_0{}^{+l})$	$P(\mathrm{fld}_2 \mid c_0{}^{+l}) = 1 - P(\mathrm{fld}_1 \mid c_0{}^{+l})$

compared to *saccade-only* ($p < .001$) and *recency-only* ($p < .001$), and more biased responses in the *location-only* condition compared to *saccades-only* ($p < .001$) and *recency-only* condition ($p < .001$). No other comparison reached significance (all p's > .473). Face and condition interacted ($F(3, 149) = 6.674$, $p < .001$, $\eta_p^2 = 0.118$). Tukey post hoc comparisons of a single-factor ANOVA on the differences between response biases (proportion of second face minus first face responses) revealed larger differences in the *saccade-only* compared to the *location-only* ($p < .001$) and the *recency-only* condition ($p = .023$). No other comparison reached significance (all p's > .113).

3.2. Logistic regression model

To predict whether participants selected the biased or unbiased face for the morph, we considered the logistic function

$$\rho(x) = 1/(1 + e^{-x})$$

and fitted the three parameters (r, s, l) in a way that the conditional probabilities given in Table 1B match the empirical means in the various conditions. The average error of the logistic regression model was 8.1% (see Appendix B). Parameter estimates indicate that oculomotor information did not play a role ($r = 0.65$, $s = -0.02$, $l = 0.54$; see Fig. 2). Together, these predictors were reliable in distinguishing between selecting the biased or unbiased face for the morph.

The results of the logistic regression analysis suggest that morph perception was biased by spatial (location) and temporal (recency) but not oculomotor (saccade) information, that is, $s \approx 0$. However, with the choice probabilities as defined by the logistic regression, the same predictions result for the *recency-only* and the *saccade-only* condition (because $\sigma(-r) \approx \sigma(-r + s)$, if $s \approx 0$, see Table 1). However, other parameterizations of the choice probabilities as a function of the three parameters (r, s, l) are conceivable. Below we develop a computational model that explains decision-making based on Luce's axiom of choice.

3.3. Computational model

Compared to the logistic regression analysis, the following computational model yields a more accurate description of the empirical data using the same three context variables saccade (s) and location (l) and recency (r) that were systematically modified in the four experiments. As for the logistic regression model, we determined for each condition the evidence I_1 and I_2 for judging the morph as resembling more Face 1 or Face 2 by a linear combination of these parameters, $I_i = 0.5 \pm r \pm s \pm l$. Because the morph contains exactly 50% of information of each face stimulus, the probability of selecting either face in the absence of experimental biases is 0.5. But as the two faces were presented after each other, the temporal order introduced a recency bias that favored second face responses. In our model we accounted for this bias by adding r in the evidence (I_2) for

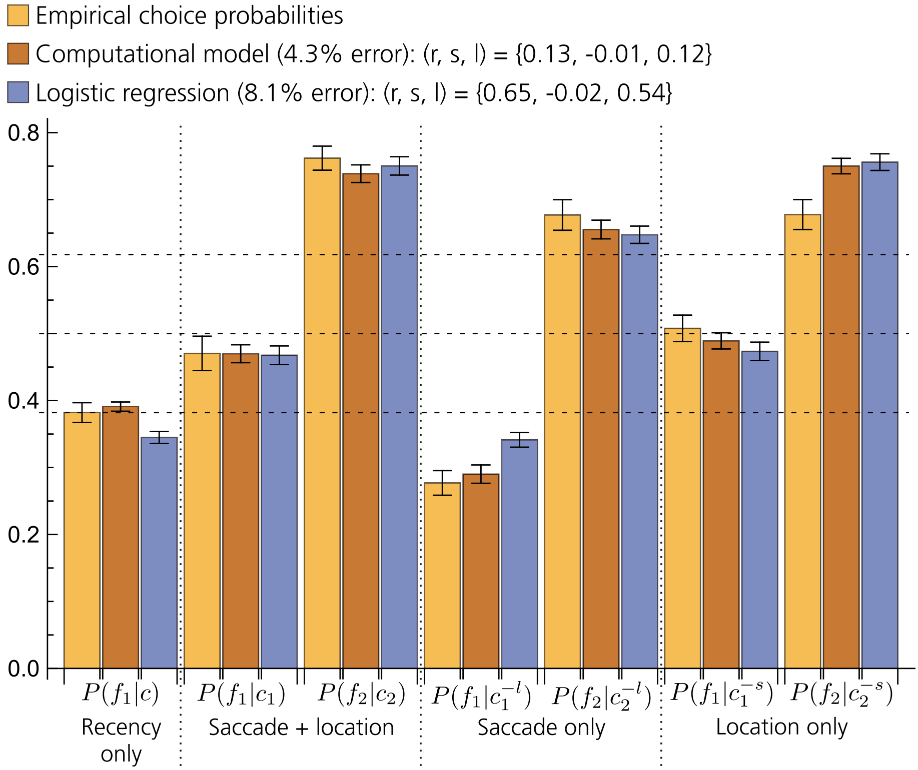

Fig. 2. Decision probabilities for the human subjects and their best fits by the computational model and the logistic regression. The seven conditional probabilities as a function of the recency, saccade, and location parameters (r, s, l) are defined in Table 1C. The dashed lines reflect chance level (0.5) and the recency effect as quantified by the *recency-only* condition. The relative error in fitting the seven choice probabilities is 4.3% for the computational model and 8.1% for the logistic regression. In addition, the computational model explains 72% of the variance in individual decision probabilities across the human subjects (see also Appendix B).

the recent Face 2 and subtracting r in the evidence (I_1) for the remote Face 1. Similarly, saccadic and location cues additionally biased the probabilities depending on a match or mismatch between face encoding and morph decoding. This was taken into account by adding or subtracting the context variable (s and/or l) in the evidence (I_i), depending on the presence or absence of the saccade or location cue, respectively.

In conditions involving spatial and/or oculomotor variations, the morph was presented in either the context (saccade and/or location) of the first or second face. In the *saccade plus location* condition, the context (saccade and location) was either the same as or complementary to the first or second face stimulus, yielding terms $+s+l$ and $-s-l$, respectively, in the conditional probabilities given in Table 1C. In the *saccade-only* condition, the context only allowed the saccade to vary, whereas the location was always new (yielding terms $+s-l$ and $-s-l$). In contrast, in the *location-only* condition, no saccades were performed before perceiving the morph, but the location varied (yielding terms $-s+l$ and $-s-l$). Note that because during recall saccades that have been enacted during encoding were missing in the location-only condition, we subtracted the saccade term ($-s$) to obtain the choice probabilities. Finally, in the *recency-only* condition, no saccades were performed and the location was always the same (yielding a term $+l$ both for the repetition of context 1 and context 2). The logistic regression model calculates choice probabilities from the differences in the evidences, for instance, $p_1 = \sigma\ (I_1-I_2)$. In our computational model we now convert evidences into choice probabilities by Luce's choice axiom (Luce, 1959, 1977), that is, by normalizing the individual evidences by their sum, $p_{\mathrm{i}} = I_i/(I_1 + I_2)$. In Luce's axiom, the ratio between the choice probabilities is just the same as the ratio between the evidences. The experimental conditions were characterized by the respective seven conditional probabilities (Table 1C) and their counter probabilities (not shown, except for the *recency-only* condition).

For instance, in the *saccade plus location* condition the evidences for Face 1 and Face 2 are $I_1 = 0.5-r + s + l$ and $I_2 = 0.5 + r - s - l$, respectively, and the conditional probability p for identifying the morph as Face 1, fId_1, given that the morph was presented in the context of Face 1, c_1, is calculated by, $P(\mathrm{fId}_1|\mathrm{c}_1) = I_i/(I_1 + I_2) = 0.5-r + s + l$. Note that in a given context c one identifies either Face 1 or 2, and the two probabilities $P(\mathrm{fId}_1|\mathrm{c})$ and $P(\mathrm{fId}_2|\mathrm{c})$ therefore sum up to one, $P(\mathrm{fId}_1|\mathrm{c}) + P(\mathrm{fId}_2|\mathrm{c}) = 1$ (as a consequence, we only show one of these probabilities in Fig. 2). In the *saccade-only* condition, where the morph's presentation always elicited a saccade associated with either of the two faces, the morph's location was always new (central). This yields evidences $I_1 = 0.5-r + s- l$ and $I_2 = 0.5 + r-s-l$, respectively, and the conditional probability $p(\mathrm{fId}_1|\mathrm{c}_1{}^{-l})$ for identifying the morph as Face 1 given the same context calculates to $P(\mathrm{fId}_1|\mathrm{c}_1{}^{-l}) = I_i/(I_1 + I_2) = (0.5 - r + s-l)/(1-2l)$. To justify the notation of the context, note that in the *saccade-only* condition the location of the morph is always different from the individual face locations (hence the superscript $-l$), and in the *location-only* condition no saccades to the morph are executed (hence the superscript $-s$).

3.3.1. Model comparison

Although both the logistic regression and the computational model closely reproduce the average of the seven independent decision probabilities with a low error, the computational model better captures the effects of recency, location, and saccades. The computational model achieved an average error of only 4.3% as compared to the average error of 8.1% for the logistic regression model (see Appendix B).

Despite the fact that the models make predictions that are roughly in the same range in terms of average accuracy across all conditions, they substantially differ in predicting choice probabilities for the *recency-only* and the *saccade-only* conditions (Fig. 2). In the *saccade-only* condition, the morph appeared in a new, previously unoccupied location (centrally, whereas the faces were encoded peripherally). This implies that the evidence both for Face 1 and Face 2 is reduced by $-l$, $I_1 = 0.5-r + s-l$ and $I_2 = 0.5 + r + s-l$. Because the logistic regression model calculates the choice probabilities as difference in evidences, $p_1 = \sigma(I_1-I_2)$ and $p_1 = \sigma(I_2-I_1)$, the location parameter l is canceled out in the argument of the logistic function σ. This is different for Luce's choice axiom. When reducing both evidences by the same amount, Luce's choice axiom disproportionally favors the more likely option, that is, the choice option that is favored even more after both evidences were symmetrically reduced.

The same effect is also observed in the experiment: Face 2 which is favored in the *recency-only* condition is favored even more in the *saccade-only* condition, where the location of the morph is always new (Figs 1 and 2). In other words, the probability for perceiving the morph as Face 1 when context 1 is repeated is lower in the *saccade-only* condition compared to the *recency-only* condition, even though a saccadic cue could be expected to increase this probability. A negative saccade parameter s cannot explain this probability decrease as it would also decrease the choice probability for Face 2 given context 2 (which is not the case, see data in Fig. 2). As recency contributed comparably to the choice probabilities in both conditions ($-r$), it is the location parameter that critically accounts for the differences in choice probabilities. While in the *saccade-only* condition, the morph location was always new (expressed in $-l$ in formula of the neuronal model, see Table 1), in the *recency-only* condition location remained congruent between encoding and morph presentation (expressed in a $+l$). Note that in the logistic regression model no such location dependence can be expressed as the location parameter l cancels in the argument of the regression function σ (Table 1). Any non-zero value of s (positive or negative) would decrease rather than increase the difference between choosing Face 1 in context 1 and choosing Face 2 in context 2 (because $\sigma(r+s)-\sigma(-r+s) < \sigma(r)-\sigma(-r)$ for $s \neq 0$, and so the best choice for the saccade parameter is $s \approx 0$). Hence, while the neuronal model closely fits the empirical data, the logistic regression model fails.

3.3.2. Best parameter fit

The parameters that minimized the mean squared error of the choice probabilities across the four conditions were $r = 0.13 \pm 0.06$, $s = -0.01 \pm 0.05$, and $l = 0.12 \pm 0.06$ (mean ± *SD*, see Fig. 2). Interestingly, like in the logistic regression model, the best fit (see Fig. 2 for a comparison between empirical and modeled data) was achieved when the influence of the saccade was approximately zero ($s = -0.01 \pm 0.05$ *SD*). Recency ($r = 0.13 \pm 0.06$) and location ($l = 0.12 \pm 0.06$) seemed to be almost equally influential. For example, the probability in the *saccade plus location* condition for selecting the first face when the morph was presented with the saccade and in the location of face$_1$ is calculated to be 0.48 (= 0.5−0.13−0.01 + 0.12). The parameter distributions obtained by the model when fitting the choice probabilities for each of the 44 human subjects individually are shown in Fig. 3.

3.3.3. Neural network architecture

To gain insight into the possible neuronal mechanisms that produce the response biases in our decision task, we propose a model that includes the process of face identification. The network is composed of three layers (Fig. 4). The first layer encodes the face image together with the image location and a possible saccade to that location. The second layer encodes the identity of the face that is determined based on the sensory and motor inputs

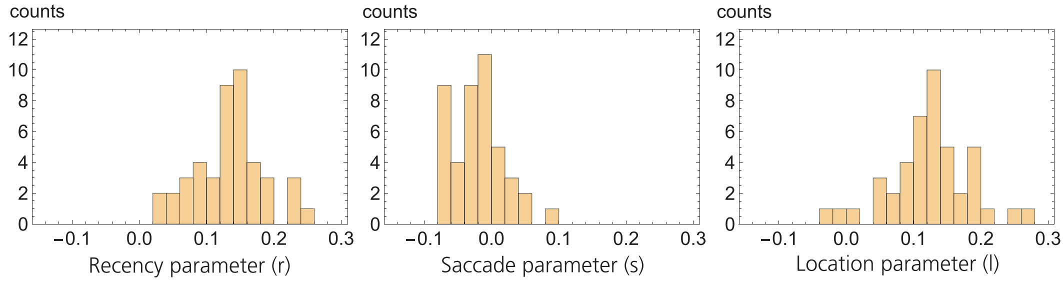

Fig. 3. Parameter distributions for the computational model (44 parameter triples [r, s, l]) that fit the choice probabilities of the 44 subjects. For each of the 44 datasets, the parameters were adapted to minimize the mean square error for the seven choice probabilities from Fig. 2. The parameter distributions reflect the interindividual differences on how recency, saccade, and location information affect the choice probabilities.

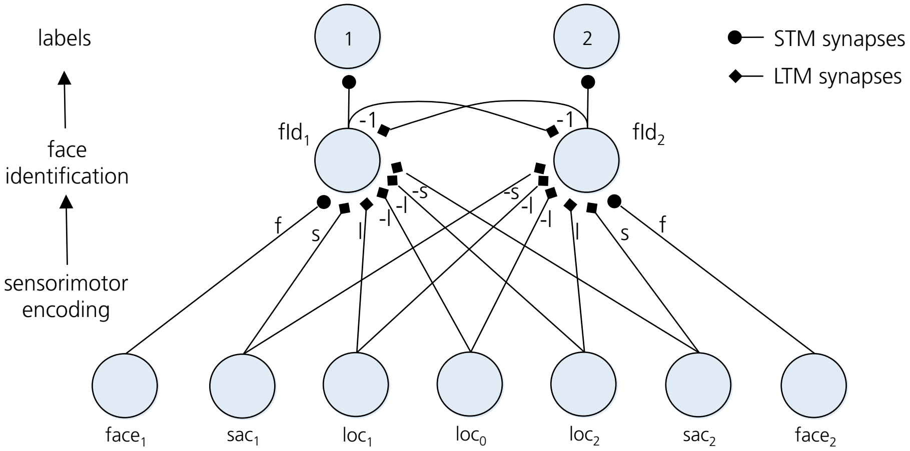

Fig. 4. Neuronal implementation. A face identification (fId) layer that acts as a winner-takes-all network receives input from a sensory-motor layer and activates one of the label units "first face" or "second face" that represent the network decisions for a specific trial. The units in the sensory-motor layer encode the face image, the saccade, and the location. Their activity (1 or 0) is transmitted through short-term memory (STM) synapses and long-term memory (LTM) synapses to the fId units (Eq. 1). The synaptic strengths of these connections are learned by a Hebbian rule.

from the first layer. The third layer encodes a label associated with the identity of a face. In our case, each face is associated with the label of being presented either as the first or second stimulus in a given experimental trial.

The first layer consists of $N = 96$ face units face$_n$ ($n = 1\ldots96$) that encode the face images, four saccade units sac$_n$ ($n = 1\ldots4$) that encode eye movements in the four diagonal directions, and five location units loc$_n$ ($n = 0\ldots4$) that encode the central and the four diagonal stimulus positions (upper left, upper right, lower left, and lower right). These units are either active (= 1) or inactive (= 0). During encoding of Face 1 or Face 2, one specific face unit becomes activated. Perceiving the morph (average of Face 1 and Face 2) then activates both of the corresponding face units. When perceiving a face image, there is at most one saccade unit (if at all) and one location unit active. The face identity layer consists of $N = 96$ mutually competing face identity units (fIds), corresponding to the N original face images used in the experiment. This second layer represents a winner-takes-all network. The total input I_j to face identity unit j is the weighted activity of the face, saccade, and location units in the first layer,

$$I_j = \sum_{n=1}^{N} f_{jn}\mathrm{face}_n + \sum_{n=1}^{4} s_{jn}\mathrm{sac}_n + \sum_{n=0}^{4} l_{jn}\mathrm{loc}_n \tag{1}$$

where f_{jn}, s_{jn}, and l_{jn} represent the synaptic strengths from the presynaptic face, saccade, and location unit ν, respectively, to the postsynaptic unit j. The face identity unit j with the strongest input, $I_j > I_i$ for $i \neq j$, is activated first, and all other units are suppressed, $\mathrm{fId}_j = 1$, whereas $\mathrm{fId}_i = 0$ for $i \neq j$, $j = 1 \ldots N$. The third layer consists of only two units that encode the sequential position of the corresponding face. Each unit represents a population of neurons.

The synaptic strengths were set as shown in Fig. 4, motivated by a Hebbian plasticity rule (see Appendix C). More specifically, we set the synaptic weight from the face_n neuron to the jth identity neuron fId_j to $f_{jn} = f$, the weight from the saccade neuron sac_n to fId_j to $s_{jn} = s$ and to fId_i $(i \neq j)$ to $s_{in} = -s$, and correspondingly, the weight from the location neuron loc_n to fId_j to $I_{jn} = l$ and to fId_i $(i \neq j)$ to $I_{in} = -l$.

3.3.4. Decision-making

When presented with a morph the two face units face_j and face_k that compose the morph are activated. The corresponding fIds receive the total input I_j and I_k and compete to become activated. We assume that a possible morph identity unit is suppressed by some external "task unit." For convenience we sequentially renumber these involved faces with 1 and 2. Based on Eq. 1 (see also Fig. 4), the total input to fId_1 and fId_2 during the presentation of the morph in a specific context has the form

$$I_1 = 0.5 - r \pm s \pm l \text{ and } I_2 = 0.5 + r \pm s \pm l \tag{2}$$

These synaptic inputs correspond to the evidences for Face 1 and Face 2 in the computational model. As before, the ± signs in I_j indicate whether the saccade and the location associated with the morph are the same or opposite as during the presentation of face j. More specifically, Eq. 2 is obtained by taking into account that at the time of the morph presentation (a) the involved face units are equally strongly activated, $\mathrm{face}_1 = \mathrm{face}_2 = 1$, (b) the connection strength from face_1 to fId_1 decayed to $\mathrm{f}_{11} = 0.5-r$, while the connection strength from face_2 to fId_2 is $\mathrm{f}_{22} = 0.5 + r$, and (c) the executed saccade and the image location can be the same (+) or different (−) as it was during the presentation of Face 1 or Face 2, respectively. The signed pair of saccade and location represents the context in which a morph image is presented, $c = (\pm s, \pm l)$, as compared to a preceding face presentation. The probability that fId_i (with i=1, 2) is activated and hence face_i is chosen when presenting the morph in context c becomes

$$P(\text{fId}_i|c) = \frac{I_i}{I_1 + I_2} \tag{3}$$

These probabilities for i = 1 and 2 represent a neuronal implementation of Luce's axiom of choice (Luce, 1959, 1977), as introduced in the computational model (s. also Table 1C). In the context of a stochastic neuronal decision-making, the synaptic drive I_i causes the neuronal population fId_i to stochastically fire with Poisson rate I_i (e.g., Churchland et al., 2011). The probability of the population fId_i firing first is given by Eq. 3, justifying Luce's axiom in neuronal terms. Once fId_i is activated, the associated label unit i gets activated and the decision for Face 1 or Face 2 is made.

4. Discussion

Based on previous findings demonstrating that memory retrieval is impaired when eye movements are restricted (e.g., Johansson & Johansson, 2014; Laeng et al., 2014), the goal of this study was to investigate distinct effects of oculomotor and spatial information on perceptual decision-making. In four conditions, participants encoded two different faces and were asked to judge whether the exact averaged image of the two faces (morph) looks more like the first or second face. Participants were more likely to perceive the morph as the face that was encoded with the same saccade or in the same location or both. However, these response biases occurred only when recent contextual information (saccade and/or location) was repeated. The *recency-only* condition confirmed that participants were more likely to perceive the morph as resembling the face that was encoded directly before the morph, thus demonstrating a recency effect. We propose a computational model explaining the response biases as a combination of memory (depending on recency), spatial, and oculomotor information. Surprisingly, predictions were best when oculomotor information was disregarded. These findings suggest that spatial rather than oculomotor information influences perceptual memory.

Previous studies demonstrated that during recall, the eyes are likely to move back to where the stimuli have been encoded (Laeng & Teodorescu, 2002; Martarelli & Mast, 2013; Spivey & Geng, 2001) and that these eye movements are functional for memory recall (Johansson & Johansson, 2014; Johansson et al., 2012; Laeng & Teodorescu, 2002). However, impairments in memory performance due to gaze manipulation in terms of the instruction (i.e., to fixate a cross vs. free viewing or to look only in an area congruent vs. incongruent to the encoding location) can be explained by a spatial mismatch between encoding and recall rather than by oculomotor mechanisms.

The computational model we suggest (that incorporates Luce's axiom of choice) has the advantage of differentially accounting for contributions of spatial and oculomotor information, providing insight into the effects of spatial location and eye movements on perceptual memory. Moreover, the suggested neural network architecture provides a plausible explanation of how these factors influence memory retrieval. This network implements the idea that visual, spatial, semantic (Ferreira et al., 2008; Richardson et al., 2009), and potentially also oculomotor information (Laeng & Teodorescu, 2002) are integrated in a single representation. It also shows that data modeling that is guided by a possible neuronal implementation can more faithfully capture the dependency of decision-making on context variables than a standard statistical approach may do. In fact, compared to the logistic regression model, reconstruction errors for the averaged choice probabilities were only half the size in the computational model.

Inference statistics (ANOVA) imply that perceptual memory is biased by location and/or saccades only when the morph contains contextual information (saccade and/or location) of the face that was encoded most recently. This can lead to the erroneous conclusion that the results simply reflect a recency effect, that is, participants were generally more likely to perceive the morph as the face that was encoded most recently, independent of spatial and oculomotor information (*recency-only* condition). Indeed, recency effects are not surprising given the compelling evidence showing that memory is best for the object that has been fixated last (Hollingworth & Henderson, 2002; Irwin & Zelinsky, 2002; Zelinsky & Loschky, 2005). The computational model can explain both effects of intervening saccades and spatial information as well as remote information (first face) by subtracting these influences from the corresponding response biases.

Explaining data with a model that can be implemented in neuronal terms is constraining. Forcing a model to take a specific form typically reduces the degrees of freedom and results in a worse fit of the data. Interestingly, this is not the case in our connectionist approach. Although the decision probability derived from our neuronal model (Table 1C) is functionally similar to the logistic regression model, our model is able to explain the different outcomes in the *saccade-only* and *recency-only* condition where the logistic regression model fails. Hence, the neuronal model captures variance in the cognitive processes that is neglected by the regression model.

In addition to the better fit, the computational model and its implementation in a neural network foster new ideas in the research field of eye movements during memory recall. For instance, the suggested neural network implementation postulates specific mechanisms of how face identities and the contextual information are encoded and reactivated during perceptual memory. Whereas the recency effect originates from short-term facilitation of synapses encoding the face identities, context information is encoded in long-term synaptic changes. Because long-term potentiation (but not short-term facilitation) accumulates across repetitions, the neuronal model predicts that repeating the identical experiment after 1 h would enhance the contextual modulation but not the recency effect.

Although our computational model achieved a better fit when oculomotor information was ignored, this does not necessarily contradict studies suggesting a functional involvement of eye movements in memory recall (Johansson & Johansson, 2014; Johansson et al., 2012; Laeng & Teodorescu, 2002; Laeng et al., 2014). One might interpret previous findings of “looking back to nothing” during recall as a consequence of spatial mismatch during encoding and recall. This interpretation is consistent with our finding that spatial information contributes to response biases in a perceptual judgment task while saccades had no influence. A similar dissociation has been found in a recent study examining the role of scanpath similarity between encoding and recognition on memory performance (Foulsham & Kingstone, 2013). Although participants looked back to regions already fixated during encoding, the temporal order (i.e., the path of fixations) was not necessarily the same. Moreover, performance was the same when scanpaths were exactly reproduced and when the same regions were refixated in a different temporal order.

The spatial component of episodic memory is emphasized in traditional theories defining episodic memory as “what-where-when” associations (e.g., Tulving, 1972). Making use of spatial information as memory cue is a well-known memory strategy (e.g., method of loci). Furthermore, memory recall is better when accomplished in the same context as during encoding (Godden & Baddley, 1975). The memory supportive effect of location or context cues might be explained by the encoding specificity principle (Tulving & Thomson, 1973), claiming that recall is best when cues that were present during encoding are also available during recall. In line with these memory theories, it has been proposed that looking back to the encoding location during recall improves performance because oculomotor cues that were enacted during encoding are reproduced during recall (Johansson & Johansson, 2014). However, the results from this study suggest that oculomotor (saccade) information does not serve as retrieval cue. One explanation for this effect could be that the brain is effectively storing the spatial context in a body- or geocentric coordinate system and that saccades in such coordinate systems are typically not correlated with object representation and are therefore suppressed. This explanation is consistent with previous research suggesting that “looking back to nothing” benefits recall of spatial object relations (Johansson & Johansson, 2014; see also Olsen, Chiew, Buchsbaum, & Ryan, 2014). Moreover, executing a saccade during encoding but maintaining fixation during recall has been found to impair memory recall (Johansson & Johansson, 2014; Johansson et al., 2012; Laeng & Teodorescu, 2002; Laeng et al., 2014). Future research will need to differentiate between allocentric (relative positions in world-fixed coordinates) and egocentric (eye positions relative to head-fixed coordinates) contributions to memory cueing.

The spatial cues in our perceptual memory task are so strong that they can almost compensate for the fading effect of memory. The present experimental data and the proposed neuronal model show that it is spatial rather than oculomotor (saccadic) information that contributes to the memory process. These findings are nevertheless in line with previous research demonstrating that visual recall is impaired when eye movements are restricted (e.g., Johansson & Johansson, 2014; Laeng et al., 2014). However, contrary to the hypothesis that this impairment stems from impeding the reactivation of a stored oculomotor trace (cf., Ferreira et al., 2008; Richardson et al., 2009; see also Laeng & Teodorescu, 2002), our model findings suggest that such memory impairments are rather caused by a spatial mismatch between encoding and recall.

Acknowledgments

We would like to thank Andrea Christinat, Mirjam Dönni, Selina Schär, Corina Schöne, Antje Stahnke, and Tammy Timmers for helping with data collection and Sergej Wüthrich for technical support.

References

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, *57*, 1060–1073. doi:10.1037/0003-066X.57.12.1060

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, *9*(1), 27–38. doi:10.1162/jocn.1997.9.1.27

Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information. *Proceedings. Biological Sciences/The Royal Society*, *259* (1355), 137–143. doi:10.1098/rspb.1995.0021

Churchland, A. K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., & Shadlen, M. N. (2011). Variance as a signature of neural computations during decision making. *Neuron*, *69*, 818–831. doi:10.1016/j.neuron.2010.12.037

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *12*(11), 405–410. doi:10.1016/j.tics.2008.07.007.

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal of Experimental Psychology: General*, *142*(1), 41–56. doi:10.1037/a0028227

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*(2), 6. doi:10.1167/8.2.6

Godden, D. R., & Baddley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*(3), 325–331. doi:10.1111/j.2044-8295.1975.tb01468.x

Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *Journal of Vision*, *10*(2), 8. doi:10.1167/10.2.8

Hebb, D. O. (1968). Concerning imagery. *Psychological Review*, *75*(6), 466. doi:10.1037/h0026771

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 113–136. doi:10.1037//0096-1523.28.1.113

Hosmer, D. W., Hosmer, T., Le Cessie, S., & Lemeshow, S. A. (1997). Comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*, *16*(9), 965–980. doi: 10.1002/(SICI)1097-0258(19970515)16:9<965:AID-SIM509>3.0.CO;2-O

Humphrey, K., & Underwood, G. (2008). Fixation sequences in imagery and in recognition during the processing of pictures of real-world scenes. *Journal of Eye Movement Research*, *2*(2), 1–15. doi:10.16910/jemr.2.2.3

Irwin, D. E., & Zelinsky, G. J. (2002). Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics*, *64*(6), 882–895. doi:10.3758/BF03196793

Johansson, R., Holsanova, J., Dewhurst, R., & Holmqvist, K. (2012). Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1289–1314. doi:10.1037/a0026585

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, *25*(1), 236–242. doi:10.1177/0956797613498260

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, doi:10.1016/j.cognition.2014.01.003

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. doi:10.1207/s15516709cog2602_3

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, *15*, 215–233. doi:10.1016/0022-2496(77)90032-3

Martarelli, C. S., & Mast, F. W. (2013). Eye movements during long-term pictorial recall. *Psychological Research*, *77*(3), 303–309. doi:10.1007/s00426-012-0439-7

Noton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science*, *171*(3968), 308–311. doi:10.1126/science.171.3968.308

Olsen, R. K., Chiew, M., Buchsbaum, B. R., & Ryan, J. D. (2014). The relationship between delay period eye movements and visuospatial memory. *Journal of Vision*, *14*(1), 8. doi:10.1167/14.1.8

Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (2000). The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(10), 1090–1104. doi:10.1109/34.879790

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, *16*(5), 295–306. doi:10.1016/S0262-8856(97)00070-X

Richardson, D. C., Altmann, G. T. M., 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. doi:10.1016/j.tics.2009.02.006.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, *65*(4), 235–241. doi:10.1007/s004260100059

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(4), 42–50. doi:10.1109/38.946630

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

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

Underwood, G., Foulsham, T., & Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: Eye movements during encoding and recognition. *Visual Cognition*, *17*(6–7), 812–834. doi:10.1080/13506280902771278

Zelinsky, G. J., & Loschky, L. C. (2005). Eye movements serialize memory for objects in scenes. *Perception & Psychophysics*, *67*(4), 676–690. doi:10.3758/BF03193524

Appendix A: Eye-tracking control data

To ensure that participants did indeed fixate fixation crosses, we used eye tracking in a pseudo random version of the *saccade plus location* condition and tested three participants (one woman, $M_{\text{age}} = 27.33$ years, $SD = 4.16$, all right handed). The stimulus material and procedure was identical to the *saccade plus location* condition. The experiment was programmed in ExperimentCenter (SensoMotoric Instruments, Teltow, Germany). Throughout the experiment, eye movements were recorded using an SMI RED system (SensoMotoric Instruments) sampling with 50 Hz, with a spatial resolution of 0.1° and gaze position accuracy of 0.5°. Calibration deviations were lower than 0.8° in all cases. The behavioral data of the three participants are similar to the group means in the *saccade plus location* condition: The mean proportion of judging the morph as resembling

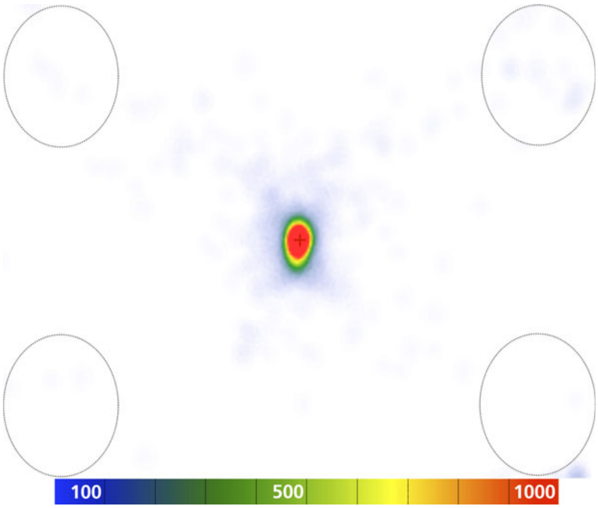

Fig. A1. Heat map during fixation. The spatial distribution of fixations is color coded for accumulated time (ms) participants spent looking at different areas during fixation crosses ranging from 1 ms (blue) to 1,000 ms (red). The maximum is 1,000 ms as this was the presentation duration of fixation cross slides. Note that ovals indicate possible face locations.

Face 1 when it was presented with the saccade and in the location of the first face was $M = 0.57$ ($SD = 0.34$), and $M = 0.75$ ($SD = 0.08$) for judging the morph as resembling the second face when it was presented in the context of Face 2. Fig. A1 confirms that participants complied with the instruction to fixate the central fixation cross.

Appendix B: Model fitting details and comparison

B.1. Details on the model fitting

To capture the variation in the data by the model, we constructed 44 model subjects. For each subject we determined a parameter triplet (r, s, l). Given such a parameter triplet, the seven independent probabilities (displayed in Fig. 2 and Table 1) can be calculated. We randomly chose 44 sets of seven empirical data points. In those conditions where <44 datasets were available, the sample sizes were completed ("bootstrapped") by doubling an appropriate number of randomly selected subsamples from other participants who performed that experiment. For each of these 44 sample sets, the mean square error on the seven data points was minimized across the parameter triplet. Mean probabilities and standard errors of the mean were calculated for the 7 probabilities across the 44 model subjects. The reported values for r, s, and l correspond to the mean of the 44 parameter triples that fit best the 44 sample sets (Fig. 2).

B.2. Details on the model comparison

To systematically compare the two models, we calculated for each model the relative deviations in the seven model mean probabilities from the seven empirical mean probabilities shown in Fig. 2, that is, $|p_{\text{model}} - p_{\text{empir}}|/p_{\text{empir}}$. These relative errors were then averaged across the seven conditions. The relative errors were calculated for 100 runs with different initial conditions for (r, s, l) and different bootstrap samplings as explained above. The histograms of these relative errors for both models show that the computational model systematically achieves half the error than the logistic regression model (Fig. B.1).

B.3. Goodness of fit and individual differences

The relative errors in estimating the decision probabilities as reported in Fig. B.1 is a simple way to quantify the model fits. In general, characterizing a goodness of fit for a binary classification is tricky and can fail in various ways (Hosmer, Hosmer, Le Cessie, & Lemeshow, 1997). But fitting seven independent values by effectively two parameters (r and l, with $s = 0$) with a relative error of 4.3% (Fig. B.1) would pass any reasonable goodness-of-fit test. Moreover, the stochasticity in the model with fixed parameter triple (r, s, l) does not fully account for the large interindividual variations among the human subjects. In fact, when considering 44 model subjects characterized by the same

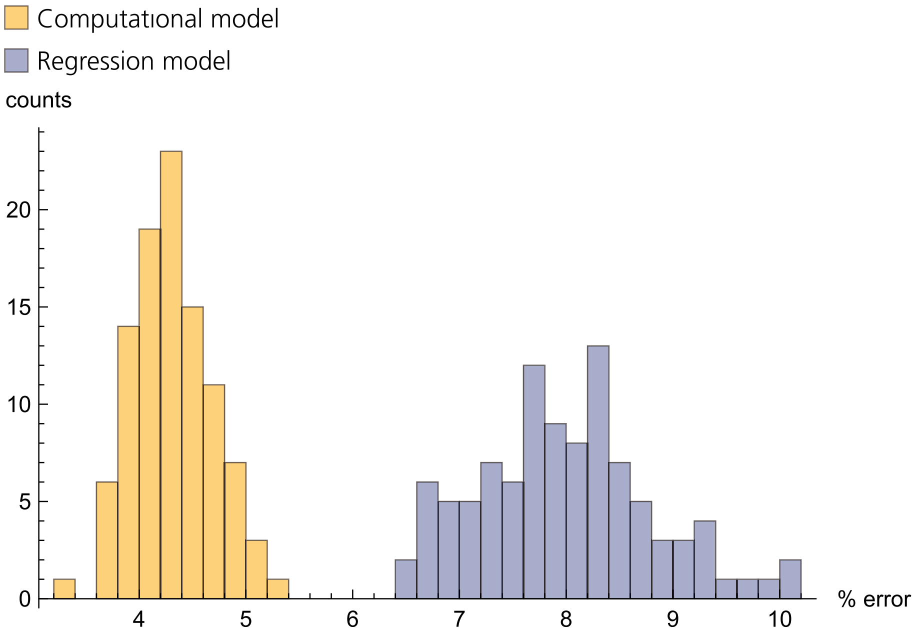

Fig. B.1. Model comparison. Histograms of the averaged relative errors for the computational model (mean error 4.3%) and the regression model (mean error 8.1%).

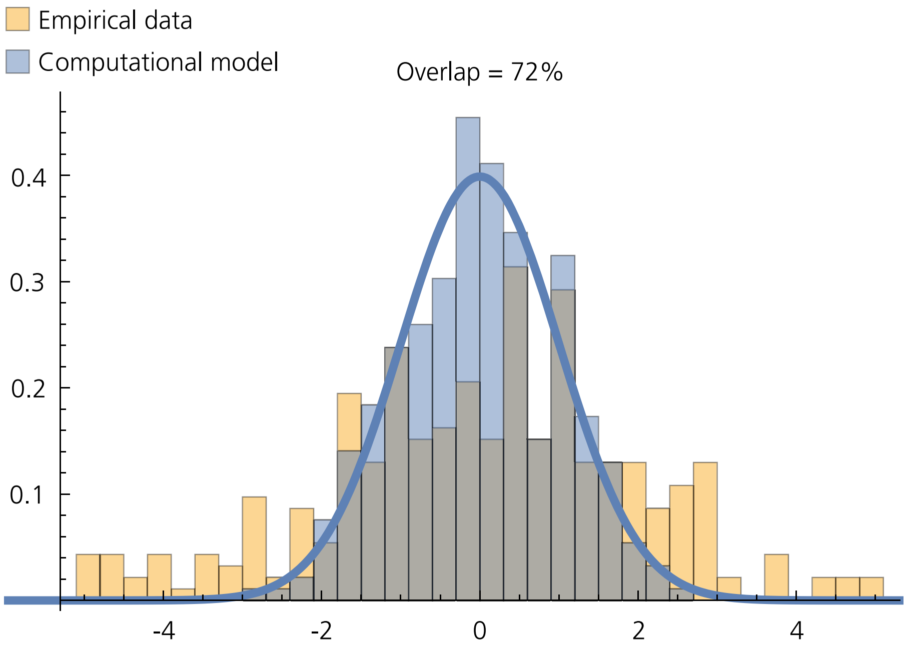

Fig. B.2. Histograms for the centered and normalized relative choice frequencies r for the 7 experimental conditions across the 44 human subjects (yellow) and model subjects (blue). The 72% histogram overlap corresponds to the "explained variance." The larger variance in the data histogram shows that the choice frequencies of the 44 human subjects are more widely distributed than the choice frequencies of a model subject that is duplicated to 44 sets. For large n, the model histogram converges to the standard normal distribution (blue curve).

parameter triplet (r, s, l), with each generating $n = 48$ binary decisions in the 7 experimental conditions, then the relative frequencies vary less than the relative frequencies vary among 44 human subjects (Fig. B.2). The centered and normalized relative choice frequencies r are defined by $r = (O - p)/\sqrt{p(1-p)/n}$. For human subjects, O represents the relative counts for one alternative in $n = 48$ trials in one of the seven experimental conditions, and p represents the average count across all 44 subjects. For the model subjects, O represents the relative count for one alternative sampled in $n = 48$ trials with probability p. For the model the frequencies r converge with large trial numbers n to a normal distribution (blue line in Fig. B2). The overlap (72%) between the histograms represents the percentage of variance in the data that is explained by the model.

Appendix C: Synapse model

It is interesting to note that the synaptic strengths from the first to the second layer can be regarded as emerging from a Hebbian plasticity rule. A synaptic weight is increased if both the pre- and postsynaptic units are active, and it is decreased or stays unaffected if the presynaptic unit is active but the postsynaptic unit is inactive (or vice versa). Because each face image is used only once for each subject, the synaptic strength f_{jn} from face unit n to face identity unit j is proportional to the product of the post- and presynaptic activity at the first face presentation, $f_{jn} \propto \mathrm{fId}_j\ \mathrm{face}_\mathrm{n}$. As for this presentation only face unit j is co-active with fId unit j, this synaptic weight is potentiated, $f_{jj} > 0$, whereas the others remain zero, $f_{jn} = 0$ for $n \neq j$. To take into account a short-term memory effect we additionally assume that the synaptic strengths f_{jj} reach a value $f_{jj} = 0.5 + r$ immediately after potentiation, where r represents a recency parameter. Formally the strength f_{jn} of the synapse from face unit n to face identity unit j is set to the product of the recency factor times the post- and presynaptic activities,

$$f_{jn} = (0.5 + r)\mathrm{fId}_j\mathrm{face}_n \tag{C.4}$$

The strength of these synapses is assumed to partially decay on a time scale of seconds such that at the time of seeing the morph, the synaptic strength from the firstly presented face unit to its corresponding fId is $f_{jj} = 0.5 - r$, whereas the strength from the secondly presented face unit to its fId is $f_{kk} = 0.5 + r$.

The synaptic strengths from saccade unit n to fId j is set when executing ($\mathrm{sac}_n = 1$) or not executing ($\mathrm{sac}_n = 0$) saccade n. Following the Hebbian rule, we set

$$s_{jn} = \begin{cases} s & \text{if } \mathrm{fId}_j = \mathrm{sac}_n = 1 \\ -s & \text{if } \mathrm{fId}_j \neq \mathrm{sac}_n \\ 0 & \text{if } \mathrm{fId}_j = \mathrm{sac}_n = 0 \end{cases} \tag{C.5}$$

with some saccade parameter s. Similarly, the synaptic strength from location unit n to decision unit j is set upon face presentation to

$$l_{jn} = \begin{cases} l & \text{if } \mathrm{fId}_j = \mathrm{loc}_n = 1 \\ -l & \text{if } \mathrm{fId}_j \neq \mathrm{loc}_n \\ 0 & \text{if } \mathrm{fId}_j = \mathrm{loc}_n = 0 \end{cases} \tag{C.6}$$

with l representing a location parameter. The free parameters of the model are (r, s, l) that encode the effect of the recency, saccade, and location.