

Memory for scenes: Refixations reflect retrieval

LINUS HOLM AND TIMO MÄNTYLÄ
University of Umeå, Umeå, Sweden

Most conceptions of episodic memory hold that reinstatement of encoding operations is essential for retrieval success, but the specific mechanisms of retrieval reinstatement are not well understood. In three experiments, we used saccadic eye movements as a window for examining reinstatement in scene recognition. In Experiment 1, participants viewed complex scenes, while number of study fixations was controlled by using a gaze-contingent paradigm. In Experiment 2, effects of stimulus saliency were minimized by directing participants' eye movements during study. At test, participants made remember/know judgments for each recognized stimulus scene. Both experiments showed that remember responses were associated with more consistent study-test fixations than false rejections (Experiments 1 and 2) and know responses (Experiment 2). In Experiment 3, we examined the causal role of gaze consistency on retrieval by manipulating participants' expectations during recognition. After studying name and scene pairs, each test scene was preceded by the same or different name as during study. Participants made more consistent eye movements following a matching, rather than mismatching, scene name. Taken together, these findings suggest that explicit recollection is a function of perceptual reconstruction and that event memory influences gaze control in this active reconstruction process.

Several lines of evidence from episodic memory research support the view that remembering is facilitated by similarities between encoding and retrieval occasions. For instance, the encoding specificity principle (Tulving & Thomson, 1973) states that memory is improved when contextual information available at encoding is also available at retrieval. Similarly, according to the transfer appropriate processing account, memory performance is improved when processes used during study are captured by the task used at test (Blaxton, 1989; Kolers, 1973; Roediger, Weldon, & Challis, 1989). For example, Kolers (1973) emphasized the importance of overlap between encoding and retrieval operations: "Recognition is achieved by virtue of the correlation between the operations carried out on two encounters with a stimulus event. The more similar the operations, the readier the recognition" (Kolers, 1973, p. 348).

Although reinstatement of encoding and retrieval operations is a general principle of episodic memory, the underlying mechanisms of reinstatement are not explicitly (or implicitly) formulated in these general views on optimal episodic memory retrieval. The empirical support for these principles is extensive (Godden & Baddeley, 1975; Morris, Bransford & Franks, 1977; Tulving, 1983; see also Smith, 1988), but neither the empirical studies nor the underlying theoretical frameworks inform us about *how* information is reinstated during the course of recognition performance. Specifically, neither the encoding specificity principle nor the transfer appropriate processing account indicate whether reinstatement is an incidental effect of the agent being in the same (vs. different) environment or whether reinstatement is driven by the agent's reconstruc-

tive retrieval process. That is, whether the agent actively seeks to reinstate encoding conditions during retrieval, or reinstates those conditions incidentally by perceiving the environment in the same way.

Jacoby and Craik (1979) proposed a more specific mechanism of retrieval reinstatement. They suggested that when the stimulus itself is insufficient to provide direct access to a representation, the agent undertakes a more elaborated reconstructive process: "We assume that such further processing is a 'bootstrapping' operation with the creation of very general, plausible contexts occurring first; if one such general context is associated with an increase in recognition familiarity, then the reconstructive operations will be refined in this direction until either full recognition occurs or the reconstructive efforts lead to no further increase in familiarity" (Jacoby & Craik, 1979, p. 7). In this view, the agent actively reinstates the memory event in a reconstructive manner, guided by information acquired over the course of the recognition task.

The primary aim of this study was to elucidate whether explicit recognition memory is characterized by active reinstatement or is better described as an incidental effect of returning to the same environment. If the agent's goal is to actively increase study-test compatibility in order to facilitate retrieval (see Jacoby & Craik, 1979), then these internal processes might have overt behavioral indicators. It is reasonable to assume that an online registration of reinstatement behavior would inform us about how retrieval is achieved in a recognition task. A general difficulty with this notion is that the behavioral indicators of reinstatement are not easily defined. However, systematic analyses of eye

L. Holm, linus.holm@psy.umu.se

movements in the context of scene recognition might provide an avenue for examining the reconstructive process of reinstatement. Specifically, the consistency of eye fixation distributions might constitute one potential index of the processes underlying retrieval reinstatement (Mäntylä & Holm, 2006), with similar patterns of fixations during study and recognition test reflecting successful retrieval.

As mentioned earlier, the encoding specificity principle and the transfer appropriate processing account are general principles of memory (Blaxton, 1989; Kolers, 1973; Roediger et al., 1989; Tulving & Thomson, 1973), but few studies have acknowledged the implications of those principles by comparing encoding and retrieval operations in terms of overt behavioral measures. One departure from this state of affairs within visual recognition memory was presented by Noton and Stark (1971). Following the ideas of Hebb (1949, 1968), they claimed that viewing a scene elicits a specific scanpath, and that recognition is characterized by the sequential reinstatement of the scanpath in that eye movements tend to follow the same pattern established during the initial viewing of the stimulus scene. Noton and Stark did not present any objective measures in support of their claim, but based their conclusions on a descriptive analysis of the scan patterns produced by their participants.

Several attempts to replicate their findings with more stringent methods have failed (Fisher, Karsh, Breitenbach, & Barnette, 1983; Locher & Nodine, 1974; Walker-Smith, Gale, & Findlay, 1977). Indeed, findings by Melcher and Kowler (2001) suggest that increased scene familiarity has no consequences on eye movements. Their participants were shown several scenes repeated over blocks, and they were instructed to report the objects presented in the scenes immediately after each trial. Despite increased scene content memory (i.e., number of objects recalled), participants were as likely to fixate new objects as objects viewed during earlier presentations. According to Melcher and Kowler, eye movements were best described as a random walk between salient areas (i.e., objects) upon subsequent scene presentations in their study.

Contrary to Melcher and Kowler, several studies of visuospatial memory (Brandt & Stark, 1997; Laeng & Teodorescu, 2002; Spivey & Geng, 2001) have shown that eye movements during encoding are reinstated during free recall, indicating that eye movements are represented together with their perceptual consequence. Furthermore, Laeng and Teodorescu (2002) found that the reinstatement of prior scan patterns was beneficial for immediate free recall of picture content. However, these studies involved very simple stimuli consisting of checkerboard configurations or single objects on uniform backgrounds. Such stimuli might have optimized conditions for eye movement reinstatement. Therefore, the generality of these findings to everyday scene recognition is rather limited.

A more flexible way of representing scenes would be by content and content location, rather than content location and content (i.e., fixation) order. Specifically, complex scenes might be represented in terms of gist and associated details (Bar, 2004; Bar et al., 2006; Brockmole, Castelano, & Henderson, 2006; Brockmole & Henderson, 2006; Davenport & Potter, 2004; Hollingworth, 2006;

Loftus, 1972, 1976, 1981). In that case, scene representations would be based on the perceptual consequence of eye movements (i.e., the distribution of fixations) but not on the sequential order of those fixations, or eye movements per se. Therefore, reinstatement of content (rather than the order of fixations; i.e., the scan path) should constitute a sufficient index of memory performance in scene recognition (see also Mäntylä & Holm, 2006).

A similar idea was implemented in two visual recognition models by Deco and Schürmann (2000a, 2000b). Their models involved an iterative hypothesis testing cycle deploying attention to diagnostic visual regions with continuous evaluation against a memory trace. It is conceivable that such a process would involve a control of eye movements (Deco & Schürmann, 2000b). Although this view seems like a fruitful start in theoretical development of episodic memory, there is little direct evidence for the suggestion to date.

Following this line of reasoning, we hypothesized that reinstatement in visual recognition is guided by previously encoded event information, and that episodic event information contributes to reconstructive retrieval operations by guiding eye movements toward diagnostic regions in a hypothesis testing manner. Therefore, episodic memory should *influence* gaze control, or “the process of directing fixation through a scene in real time in the service of ongoing perceptual, cognitive and behavioral activity” (Henderson, 2003, p. 498).

A possible reason for the limited effects of eye movements on memory performance in past research (see, e.g., Fisher et al., 1983) is that episodic memory has been assessed in terms of overall recognition performance, rather than in more specific components such as levels of confidence or recollection and familiarity. Several lines of evidence indicate that recognition memory reflects two distinct states of awareness (Jacoby, 1991; Mandler, 1980; Tulving, 1983; see also Yonelinas, 2002, for a review). One line of evidence for this dual-component view of recognition comes from studies that have adopted an *experiential approach* to memory and awareness (Tulving, 1985; see also Gardiner & Java, 1993; Rajaram & Roediger, 1996, for reviews). According to this approach, an event is recognized when its occurrence brings to mind some specific experience in which the event was originally encoded (“remembering”). Alternatively, an event is recognized, not because of specific images or experiences, but because of feelings of familiarity that can be attributed to it (“knowing”). That is, the individual is able to tell that a given item is familiar, but does not remember actually seeing the item, or have a conscious recollection of it.

The underlying mechanisms of remember and know judgments are yet unclear (see, e.g., Wixted, 2007, for a discussion). However, this controversy relates to the explanations of remember and know judgments in terms of recollection and familiarity, respectively, not to whether a division of recognition judgments based on two qualitatively different experiences is subjectively meaningful. In addition, remember responses generally reflect the output of a strong memory (Wixted 2007; Yonelinas, 2002) and should constitute a more sensitive measure of episodic or source memory than know responses, which are assumed to reflect

the output of weaker source memory. Therefore, we reasoned that remember (rather than know) responses would be systematically related to consistency of eye movements. This hypothesis was based on the notions that explicit recollection (in terms of remember judgments and source recall; see Johnson, Hashtroudi, & Lindsay, 1993, for a review) reflects encoding of distinctive event attributes (Jacoby & Dallas, 1981; Johnson et al., 1993; Mäntylä, 1997; Rajaram, 1996), and that encoding of specific details requires focal attention (i.e., fixations, see Hollingworth & Henderson, 2002). Furthermore, perceptual congruency between study and test has been shown to affect the incidence of remember, rather than know, responses (Rajaram, 1996).

Mäntylä and Holm (2006) provided partial support for the hypothesis that consistency of fixations differentiates recollective experiences in the context of face recognition. In their study, participants indicated their recognition experience by making remember/know judgments. Eye movement consistency was defined in terms of proportion of test fixations landing within approximately two degrees of *any* study fixation of a stimulus scene. This dependent measure should be adequate, considering that the general recognition task involves differentiating new items from old ones. Therefore, similarity in the overall distribution of fixations between study and test might not be crucial, as long as the limited set of reinstated details provide sufficient discriminative information to identify the scene (see also Brainerd, Reyna, & Mojardin, 1999; Loftus, 1976). Furthermore, if higher cognitive functions, such as retrieval from episodic memory, influence gaze control, one might not expect this influence to work on a tight fixation to fixation basis, because these processes can be expected to operate on different time scales (Ballard, Hayhoe, Pook, & Rao, 1997). It would be more reasonable to expect refixations to be interleaved by fixations of new regions even for subsequently recognized scenes. In addition, multiple refixations of the same region during a test trial might suggest that the second refixation constitutes a confirmation of a memory hypothesis formed after the first refixation of that region. Therefore, a dependent measure assigning equal importance to any test refixation should be favorable over say an intuitively reasonable measure of number of different study regions refixated at test.

One difficulty with the consistency measure is its close relation to number of fixations. Specifically, because the area covered by fixations increases as a function of number of fixations, increasing the area covered by *study* fixations also *increases* the likelihood of a subsequent test refixation. Similarly, the refixation measure should *decrease* with number of *test* fixations, because entropy increases with the series length, and each additional test fixation is hence associated with a reduced probability of matching a study fixation. Mäntylä and Holm (2006) did not control for number of study and test fixations in their analysis of eye movement consistency.

To examine the relation between eye movements and memory performance in terms of eye movement consistency it is therefore necessary to control for number of fixations. In addition, Loftus (1972) found that scene memory was determined by number of study fixations

rather than presentation time. Therefore, controlling for number of fixations would also serve the purpose of keeping the mnemonic consequences of scene study constant.

In this study, we constrained number of study fixations by using a gaze-contingent procedure (Experiments 1 and 3) and a moving-dot paradigm (Experiment 2). In addition, we limited the analysis of test data to the first three participant selected fixations. This procedure made levels of refixations in memory judgments comparable within and between experiments.

Another difficulty with the consistency measure is that salient scene attributes might attract the viewers' attention during study and test, and produce a high degree of eye movement consistency. For instance, a perceptually outstanding feature such as a bright moon on a homogeneous night sky might influence a participant to direct his or her gaze to the moon during the two different presentations, and hence receive a high level of consistency. This kind of consistency might facilitate recognition, but gaze control could not be uniquely attributed to episodic memory (see Mäntylä & Holm, 2006).

In Experiment 2, we controlled for stimulus saliency by constraining participants' eye movements during study. Specifically, participants studied one of two sets of similarly salient scene regions indicated by a moving dot. All participants saw the same scenes, but they studied them differently, depending on which set of regions were indicated. In a later recognition test, participants were allowed to move their eyes freely. If recollection involves reprocessing of highly resolved stimulus information, regions fixated during study should be favored over *similarly* salient regions during a recognition test with free eye movements. Alternatively, if recollection is not related to perceptual reinstatement, fixations should be evenly distributed between equally salient regions.

Finally, in Experiment 3, we addressed the question of active reconstruction in scene recognition by investigating whether scene memory would influence eye movements during recognition. Specifically, if memory for a previously studied scene is causally, rather than incidentally, related to gaze control, this episodic information might contribute to reconstructive retrieval operations by guiding eye movements toward diagnostic regions of the test scene. We reasoned that, if scene recognition can be characterized by such iterative hypothesis testing (see Deco & Schürmann, 2000b), then manipulating the validity of that hypothesis should influence gaze control during recognition. Specifically, prior information about what view to expect should reactivate participants' event representations. If those episodic representations influence eye movements in a confirmatory way, one would expect a higher degree of eye movement reinstatement, compared to nonconfirmatory cases. This approach would constitute a direct test of Jacoby and Craik's (1979) "bootstrapping" notion of reconstructive retrieval.

EXPERIMENT 1

Method

Participants. Twenty-three Umeå University undergraduates (14 male, 9 female) participated in the experiment for payment (approx-



Figure 1. An example of a stimulus picture used in all experiments. Original stimulus pictures were in color. Picture adapted and reprinted with permission of the artist, Jane Wooster Scott.

mately US\$6). They were between 19 and 27 years old, and had no prior experience of similar experiments.

Materials. The stimulus materials comprised 76 digital landscape pictures of paintings (by the artist Jane Wooster Scott) in color, including eight filler and practice items. The stimuli were selected from a Web-based picture gallery (www.woosterscott.com), which provided relatively similar paintings of landscapes (see Figure 1). The pictures were divided in two sets of 36 items. Each subset appeared equally often as targets and distractors. At test, participants were presented with 36 studied pictures intermixed with 36 nonstudied pictures. In addition, 4 practice items were shown at the beginning of the test, yielding 76 items. In the present setting, the stimulus picture corresponded to approximately 40×30 degrees of visual angle. Stimulus size conditions remained the same in Experiments 2 and 3.

Apparatus. An EyeLink I eyetracker (SR Research Ltd.) was used to monitor participants' eye movements during study and test. The sampling rate of the system was 250 Hz and the spatial accuracy was between 0.5° and 1° of the visual angle. The pictures were presented in 800×600 pixel resolution on a 19-in. computer screen with an 85 Hz refresh rate.

Procedure. Each individually tested participant was seated approximately 60 cm from a computer screen. A chinrest was used to maintain a constant viewing distance and reduce head movements. The participants were informed that a series of pictures were to be studied in preparation for a later recognition test, but the gaze-contingent nature of the experiment was not mentioned. The study phase was preceded by an eyetracker calibration procedure which lasted approximately 5 min. During study, each stimulus picture was presented for 18 fixations. Once the threshold of 18 fixations was reached, a frame consisting of a central black dot on a gray background replaced the stimulus picture. The subsequent stimulus item was presented once the participant fixated the dot. Test followed immediately after the study phase without an intervening eyetracker calibration. The experimenter informed the participants that some of the previously presented pictures would appear along with new pictures, and that they should decide whether a given test item was old or new. The participants were also instructed to describe their recollective experience by answering "remember," "know," or "guess," respectively, for each recognized item (see also Mäntylä & Holm, 2006). A remember response was described as a recollection of any thoughts associated to the time when the picture was presented. A know response was described as a feeling of familiarity without any recollection of episodic details. To reduce the dependency between the remember/know judgments, the participants were told to respond "guess" if they could not decide whether they "remem-

bered" or "knew" an item. The participants were also told to indicate their old/new judgments by pressing a button and subsequently provide their verbal responses. Once the button was pressed (which also terminated stimulus presentation), or after the deadline of four test fixations, the participants provided a self-paced remember/know response. Finally, the participants were asked whether they had experienced differences in the presentation rate. Although a few participants noted differences, none reported having discovered the connection between fixation frequency and presentation time.

Fixation detection and measurement. Fixation was defined online by a computer algorithm employing acceleration ($<800^\circ/\text{sec}^2$), velocity ($<30^\circ/\text{sec}$), and motion criteria ($<0.15^\circ$) calculated over a cycle of four samples. Fixation information was fed back to the computer supporting the stimulus presentation for use in the gaze-contingent procedure. In addition to the mentioned criteria, only fixations with durations above 100 msec were considered. We used an "end of fixation signal" to adjust the count toward the threshold number of fixations in the gaze-contingent procedure. This procedure introduced a delay from the last fixation end to stimulus offset by an average of 27 msec ($SD = 3.6$).

Results and Discussion

Memory performance. One participant was excluded due to poor tracker calibration. Overall recognition performance for the remaining participants was .64 hits and .29 false alarms (FAs), producing an average sensitivity score (d') of 0.97, $SD = 0.35$. Decomposing hits into underlying components yielded the average values of .33 and .30 for remember and know responses, respectively (due to few observations, guess responses were eliminated from subsequent analyses). The corresponding FAs were .06 and .21 for remember and know responses, respectively. Because the FA levels were fairly high, we also analyzed the data separately by excluding seven participants lacking discrimination with respect to their know responses. A prerequisite for differentiating response types with eye movement consistency measures, is that participants demonstrate reasonable memory discriminability for both response types. The remaining analyses of Experiment 1 are based on this reduced sample, but the overall pattern of results is similar for the whole sample. For this reduced sample, overall recognition performance was .61 hits and .24 FAs, and an average d' of 1.06 ($SD = 0.37$). Decomposing the overall hit data into components of recollective experience, remember responses constituted .27 and know responses .32, whereas guess responses made up .03 of hits. The corresponding FA rates were .05, .18, and .02 for remember, know, and guess responses, respectively.

Recognition and eye movements. Recognition responses with less than four test fixations (i.e., fast recognition responses) were eliminated, reducing data by 2%. Proportion of refixations was based on the number of test fixations landing on regions viewed during study.¹ The measure was calculated by summing the number of test fixations within 44 pixels (approximately 2° of visual angle) of any study fixation. The sum was then normalized by dividing with the total number of test fixations, thus yielding a relative measure of test refixations.

The first fixation would not contribute to the sensitivity of the refixation measure, because it invariably constituted a refixation, being directed toward the center of the scene. In addition, the first fixation region was not selected by

the participant, and could therefore not involve any mnemonic influence. Therefore, the first test fixation was omitted from further analysis. This procedure was employed throughout this study, the line of reasoning being similar for Experiment 2 and Experiment 3 (see also Mäntylä & Holm, 2006). Furthermore, because we were primarily interested in eye movements preceding the recognition response, only the first three test fixations (i.e., three first fixations after the initial central fixation) were used in the eye movement analyses. Because entropy increases with series length, the expected proportion of refixations would decrease with series length, and therefore comparisons between series of different lengths would be unequal. In addition, keeping the number of test fixations constant made comparison between recognition judgments within each experiment easier to interpret. This procedure also made the refixation data comparable between Experiments 1 and 3. However, an analysis of all test fixations preceding a response showed the same pattern of results in all experiments of this study.

The average proportions of refixations were .60, .54, and .51 for remember, know, and misses, respectively. These numeric differences seem rather moderate, but it should be noted that a baseline² value of refixations for nonsimilar paired study and test items was .34. Therefore, the differences between response types left to vary after controlling for a baseline suggest that relative differences among response types was fairly high. A repeated ANOVA on the refixation data yielded a significant effect of response type [$F(2,28) = 3.43$, $MS_e = .009$, $p = .046$]. The effect size of response type was $\eta^2 = .19$. Subsequent contrast tests showed that remember responses were not associated with a significantly higher level of refixations than know responses [$F(1,14) = 3.11$, $MS_e = .020$, $p = .10$] but significantly higher than misses [$F(1,14) = 7.94$, $MS_e = .015$, $p = .014$]. In turn, the difference between know responses and misses was not reliable ($F < 1$).

Taken together, the results of Experiment 1 support the notion that eye movement consistency is systematically related to the nature of recollective experience, even when numbers of study and test fixations were controlled for by using a gaze-contingent procedure. These findings are consistent with Mäntylä and Holm (2006) in that eye movement consistency differentiated recognition judgment, with explicit recollection reflecting a greater degree of overlap between study and test fixations than misses. In contrast to Mäntylä and Holm, the numeric difference in consistency between remember and know was not statistically significant. This discrepancy might reflect differences in stimulus materials and target–distractor similarity.

EXPERIMENT 2

Although the results of Experiment 1 suggested that eye movement consistency is related to scene recognition, it remains unclear to what extent this consistency could be attributed to episodic memory or idiosyncratic scene preferences consistent across scene repetitions (Mannan, Ruddock, & Wooding, 1997). Specifically, a participant might be inclined to direct his gaze toward for example, flowers in every picture, and hence increase the amount of

refixations. Such idiosyncrasies might reduce the amount of variability left to be associated with episodic memory. In Experiment 2, we attempted to increase the sensitivity of the refixation measure by limiting the influence of idiosyncratic scene preferences. To this end, we induced experimental control over fixation distribution during study by adopting the follow-the-dot paradigm of Hollingworth (2004). In this paradigm, a dot is flashed sequentially at different regions in the stimulus picture, and participants are instructed to view the scene by focusing on the regions indicated by the moving dot.

As noted earlier, number of study fixations is positively related to the likelihood of a subsequent test refixation. An implication of this is that the sensitivity of the refixation measure would be reduced with a large number of study fixations. To increase the sensitivity of the consistency measure, the number of study regions (i.e., dot locations) was reduced to five in Experiment 2. Furthermore, to increase discriminability in recognition performance, each stimulus item was repeated three times during study, and the response deadline at test was set to 5 sec.

Method

Participants. Sixteen Umeå University undergraduates (7 male, 9 female) participated in the experiment for payment (approximately US\$6). They were between 19 and 29 years old and had no prior experience of similar experiments.

Materials. The study fixation data of Experiment 1 were used to extract locations for the moving dot. First, all fixation coordinates of each stimulus picture were grouped to 10 clusters each by using a k-means cluster algorithm (see Figure 2). The minimum distance between any two cluster centers was always in excess of 5°. The clusters were then divided in two sets of five for each stimulus item. The salience of each set was assessed by comparing the proportion of fixations from Experiment 1 data within 2° of any cluster center. Two study versions were created from these cluster centers, each version consisting of items with similarly salient subsets of regions, receiving an average proportion of fixations of .21 and .22, respectively. The difference in terms of subset preference among participants of Experiment 1 was not significant [$t(21) = 1.29$, $p = .21$]. Thus the salience of the versions was fairly similar. Furthermore, a prior likelihood of about .21 indicated that the regions were rather attractive, considering that they only covered about 2% of the picture area, and should hence constitute memorable content.

The same pictures were used as in Experiment 1, and each item appeared equally often as targets and distractors. Furthermore, half of the participants studied each stimulus item by following the movement of one subset of dots, whereas the other half followed the other subset. The dots always appeared in the same order for a given item.

Procedure. Participants were instructed to memorize a series of scenes for a subsequent memory task by moving their gaze between specific scene regions. They were informed that a moving green dot would indicate these scene regions, and that they should follow the movement of the dot. Each dot was located on a cluster center. The radius of the dot was 25 pixels and appeared for 150 msec, immediately followed by a smaller dot (radius = 10 pixels) with the same center for another 150 msec. Total amount of time that the dot masked a region was thus 300 msec. It seems unlikely that this would interrupt foveal processing of the indicated region much, because saccade onset latency toward a flashed region would be expected to be around 200 msec (Deubel & Schneider, 1996). In addition, extending the duration the dot was presented beyond the expected latency would increase the chance that participants accurately targeted the dot (i.e., allowed for corrective saccades to target). Participants were instructed to keep fixating the center of the indicated region until the next dot succession appeared, 1,000 msec later. Earlier



Figure 2. A stimulus scene with an overlaid saliency map based on participants' fixations in Experiment 1. Each fixation is represented by a Gaussian light intensity distribution with a standard deviation of 1° (see also Pomplun, Ritter, & Velichovsky, 1996). The dots indicate the 10 cluster centers derived from the participants' fixations (dark dots, Set 1; light dots, Set 2). Note that only one dot was shown at a time during the experiment and participants were never presented with an overlaid saliency map. Picture adapted and reprinted with permission of the artist, Jane Wooster Scott.

pilot studies had shown that an initial large dot was easier to detect with peripheral vision, whereas the smaller subsequent dot helped to reduce fixation variability within the indicated region. The dots appeared at five different locations for each stimulus picture. Participants practiced this procedure by observing four items repeated twice before study. None of the participants reported any problems in following the dot. A frame including a green dot on a gray background intervened between every stimulus item. Participants were told to fixate the dot and initiate the presentation of the next stimulus picture by pressing a key. The dot corresponded to one of the cluster centers in the following stimulus item, hence contrary to Experiment 1, the initial position varied according to each stimulus item in Experiment 2. At test, which followed immediately after study, participants were instructed to respond in the same way as in Experiment 1, while their eye movements were registered.

Results and Discussion

Memory performance. Overall recognition performance was .68 hits and .16 FAs. The average sensitivity score (d') was 1.63 ($SD = 0.47$). Decomposed into response types the hit proportions were .47 and .21 for remember and know responses, respectively, and FA rates were .04 and .13 for remember and know responses, respectively. Due to few observations, guess responses were eliminated from subsequent analyses.

Recognition and eye movements. Recognition responses with less than four test fixations (i.e., fast recognition responses) were eliminated, reducing data by a total of two trials across all participants and trials. The refixation measure was calculated by relating each participant's test fixations to the moving dot pattern presented during study. Any test fixation within 2° of its nearest study fixation (i.e., the dot position) was considered as a refixation. As in Experiment 1, the first test fixation was removed from further analyses. Furthermore, to make the procedure more consistent with Experiment 1, only the next three test

fixations were analyzed (i.e., fixations number 2–4). The fixation data were analyzed by including both the whole set of fixations and only three test fixations (as in Experiment 1). These analyses yielded similar patterns of results, and therefore, only the latter analyses are reported here.

The average proportions of refixations were .32, .23, and .19 for remember responses, know responses and misses, respectively. The refixation data was analyzed in a repeated measures ANOVA, yielding a significant main effect of response type [$F(2,30) = 10.87$, $MS_e = .007$, $\eta^2 = .42$, $p = .001$]. Subsequent contrast tests yielded significant differences between remember and know responses [$F(1,15) = 6.68$, $MS_e = .019$, $p = .021$]. In turn, know responses were not associated with reliably higher levels of refixations than misses [$F(1,15) = 3.12$, $MS_e = .010$, $p = .10$]. The random baseline proportion of refixation was calculated in the same way as in Experiment 1. This analysis yielded an average of .14. It should be noted that the difference between false rejections and the random baseline was only 5%.

To assess the effects of scene salience, we also examined whether test fixations toward studied regions was prioritized over a baseline of similarly salient regions. This additional analysis would provide a purer measure of reinstatement. If recollection reflects reinstatement of perceptual information, one might expect a higher proportion of the saccades to be directed toward the set of studied regions over the *similarly* salient alternative set of regions. Alternatively, if recollection is not related to perceptual reinstatement, eye movements should be distributed evenly between similarly salient regions prior to recollection.

To this end, we obtained a preference index. First, we acquired the proportion of test fixations within 2° of any region from the nonstudied set of regions. Subsequently, we divided the proportion of refixations with the sum of the proportion of refixations of studied and nonstudied regions. In this index, values above .5 indicated a preference for studied over nonstudied regions. For the first three test fixations, this measure yielded an average of .57, .43, and .37 for remember, know, and misses, respectively. The sphericity assumption was not reached (Huynh-Feldt = .68), but a repeated measures ANOVA with adjusted degrees of freedom (lower bound method) produced a significant effect of response type [$F(1,15) = 7.76$, $MS_e = .047$, $p = .014$]. Individual comparisons indicated that remember responses were related to a higher preference for studied regions than were know responses [$t(15) = 2.23$, $p = .042$] and misses [$t(15) = 3.87$, $p = .002$]. The difference between know responses and misses was not significant [$t(15) = 1.47$, $p = .16$]. Importantly, separate comparisons of response types to a baseline preference index of .5 showed that remember responses yielded a significantly greater value [$t(15) = 2.79$, $p = .014$]. In addition, know responses were not associated with a reliably lower preference for studied regions [$t(15) = 1.32$, $p = .21$]. Finally, misses were associated with a significantly lower preference value than the expected .5 [$t(15) = 3.75$, $p = .002$]. In other words, only remember responses were related to an increased probability of fixating studied regions over equally salient nonstudied regions.

Taken together, the findings of Experiment 2 provide additional support for our hypothesis that recognition is related to study–test consistency in terms of eye movements. Contrary to Experiment 1, in which item saliency might have contributed to the observed effects of gaze consistency, these item-specific effects were minimized in Experiment 2 by experimentally controlling the distribution of fixations. Furthermore, Experiment 2 showed that recollection is related to a higher degree of study–test consistency than familiarity-based recognition, as indicated by a greater proportion of refixations for remember than know responses. In addition, the effect size of Experiment 2 was quite large, especially if viewed in relation to the random baseline of refixations.

EXPERIMENT 3

Experiments 1 and 2 addressed and clarified several methodological issues regarding eye movement consistency and its relation to scene memory, but they did not directly test the causal direction of that relationship. It might be the case that the eyes are incidentally directed to studied regions during test, which in turn support recollection by increasing study–test consistency. Alternatively, the observer might have an initial feeling of familiarity for the scene, and direct his or her gaze to diagnostic regions to confirm that hypothesis. Another potential problem with Experiments 1 and 2 is that the instructions for the remember response might have biased participants to base their responses on specific scene details which would hence increase the amount of refixations related to those responses. Furthermore, Experiments 1 and 2 might constitute underestimations of memory influence on eye movement consistency, because the participant is not prepared for which specific picture will be presented next, and hence initial fixations are more likely to be influenced bottom up by perceptual features. Instead, in many natural recognition situations we have expectations prior to seeing the scene, such as expecting what view will appear when we turn around a corner.

In Experiment 3, we examined gaze consistency and its causal relation to scene recognition by using an alternative strategy. As mentioned earlier, we reasoned that episodic information might contribute to reconstructive retrieval operations by guiding eye movements toward diagnostic regions of the test scene. Specifically, prior information about what scene to expect should reactivate some of the scene information encoded during the earlier study phase. If this episodic event information has a causal effect on gaze control, valid expectations should produce a higher degree of eye movement reinstatement than invalid expectations.

In Experiment 3, participants studied scenes (a subset of the paintings of landscapes used in Experiments 1 and 2) paired with an arbitrary label (e.g., “Blygtorp” [Shythorpe]). At test, half of the labels were presented before the same stimulus scene and the remaining labels with different (but previously studied) scenes. Participants were told to decide whether a given scene corresponded to its scene name, while their eye movements were tracked.

Providing the name *prior* to presentation of the scene was expected to trigger a retrieval process in which some

of the previously associated scene regions are reactivated (i.e., “Shythorpe” reinstates portions of the corresponding scene representation). This reinstatement of event information was expected to have systematic (causal) effects on subsequent eye movements during the course of recognition. Specifically, for matching scenes, one would expect the participant to refixate previously fixated regions to a higher extent, compared to the case where the scene does not correspond to the name, but is equally familiar due to its presentation during study. Alternatively, if memory does not affect eye movement distribution, there should be no difference between the conditions.

Method

Participants. Twenty-four Umeå University undergraduates (8 male, 16 female) participated in the experiment for payment (approximately US\$13). They were between 20 and 37 years old and had no prior experience of similar experiments. Participants had normal or corrected-to-normal vision (contact lenses). All participants passed the Ishihara color screen test (Ishihara, 1984).

Materials. Twenty-four stimulus scenes were selected from the same pool of landscape paintings as in Experiments 1 and 2. To mimic Swedish village names, we recombined 24 different Swedish village suffixes with 24 adjectives or nouns (e.g., “Blygtorp” [Shythorpe]). There was no obvious relationship between the village name and scene contents.

Eight different test lists of 24 labels and scenes were created. Each scene item occurred equally often as matched and mismatched across participants. Because the mismatched items might include regions of interest that largely overlapped with regions of interest in the indicated scene, four random label–scene combinations were created to reduce the risk for such item effects. Each label and scene was presented only once at test. Matched and mismatched label–scene pairs were randomly mixed.

Procedure. Participants were informed that they should learn the names of a set of village scenes, which would be presented repeatedly during a study phase. Participants were told that they had to learn the names of each picture in preparation for a later recognition test. In addition, participants were told to make judgments about their future recognition performance of each scene picture, following the presentation of the corresponding village name. Participants were told to use a four grade scale to indicate their confidence, from 0 = *certain they would not identify the corresponding scene* to 3 = *certain they would identify the corresponding scene*. Each village name was presented for 4.5 sec followed by a gray frame with a central black dot. At this stage, participants indicated their recognition confidence. Immediately after their judgment, the corresponding stimulus scene was shown. Each scene was presented for 18 fixations in the same manner as in Experiment 1. The labels and pictures were shown three times each in three blocks of randomized order.

At test, participants were shown a village label for 4.5 sec, followed by the calibration frame and subsequently a scene picture. Participants were told to decide whether the scene corresponded to the village name. Participants responded by pressing a button on a button box (which terminated scene presentation), and subsequently reported whether the picture did or did not correspond to the village name. The instructions emphasized accuracy over speed in identification judgment. Similar to Experiment 1, a response deadline of four test fixations was employed. Post experiment interviews indicated that several participants noted some variability in scene exposure duration at the study and test phase, respectively, but none of them revealed the gaze-contingent nature of the presentation.

Results and Discussion

Participants’ confidence ratings of future scene recognition memory increased significantly across the study

blocks, averaging 0.0, 1.3, and 2.1 for the first, second, and third blocks, respectively [$F(2,46) = 268.00$, $MS_e = 0.1$, $p < .001$]. These data suggest that participants' memory for names and scenes increased across study blocks. It should also be considered that the confidence ratings were made before the actual scene item was presented, hence participants confidence prior to test should have been even higher than that of Block 3. Proportion of refixations between the first and the second as well as between the second and the third block was calculated. Repeated presentation of study items was not reflected in consistency of eye movements, averaging from .61 to .60 for comparisons between Blocks 1 and 2 and Blocks 2 and 3, respectively. However, the absence of effects was expected in that participants might have attempted to cover different scene regions across repeated study trials. Consistent with the confidence data, participants were rather sensitive in identifying the scenes, and produced an average of .85 hits and .13 FAs. Nine participants did not make a single FA, and 6 participants correctly identified all targets. In order to calculate d' , these FA and hit values were replaced by .01 and .99, respectively. The average d' calculated was 2.69 ($SD = 1.29$). Fast responses (i.e., <4 test fixations) were removed from the eye movement analyses. This procedure limited the amount of data by 29%.³

The refixation measure was computed relating the eye fixation data between Block 3 and the test session. It should be noted that a comparison with the final study block constitutes the most adequate contrast, because the fixations from Block 3 represent the last visual information acquired by the participant, and should hence reflect the strongest scene content memory. These analyses indicated that valid scene expectations produce a higher degree of eye movement reinstatement than invalid expectations. Specifically, the average proportion of refixations was .69 and .58 for the matched and mismatched conditions, respectively. The difference between the conditions was statistically reliable [$t(23) = 2.34$, $p = .03$].

A random baseline for the refixation measure was calculated in the same way as in the previous experiments and showed an average of .39. Against this baseline, the relative difference in refixations between the stimulus match conditions was quite large (i.e., 77% and 49% increments in consistency for the matched and mismatched conditions, respectively). Overall, Experiment 3 supported our hypothesis that prior expectations about a scene influence gaze control by producing eye movements confirming the episodic memory of the scene.

GENERAL DISCUSSION

Most conceptions of episodic memory hold that compatibility of encoding and retrieval operations facilitates retrieval (Roediger et al., 1989; Tulving & Thomson, 1973), but remain unspecific as to the mechanisms involved in that reinstatement. The main purpose of this study was to examine some of these mechanisms by using eye movements as a window to reconstructive retrieval processes, because eye movements reflect attentional processes and focal vision provides high-acuity information.

We assumed that consistency of eye movements between study and test should be closely associated with compatibility of encoding and retrieval operations (i.e., degree of reinstatement). We hypothesized that such reinstatement should be related to recognition memory performance, and specifically characterize remember responses. Another central hypothesis of this study was that memory for a previously studied scene is causally, rather than incidentally, related to gaze control, and that this episodic information contributes to reconstructive retrieval operations by guiding eye movements toward diagnostic regions of the test scene.

Although this hypothesis seems intuitive in light of most conceptions of episodic memory, earlier research on eye movements and recognition memory provided mixed or weak support for perceptual reinstatement in recognition memory (Fisher et al., 1983; Locher & Nodine, 1974; Melcher & Kowler, 2001; Noton & Stark, 1971; Walker-Smith, Gale, & Findlay, 1977). A limited number of studies provided some support for the hypothesis (Noton & Stark, 1971; Parker, 1978) but these findings were based on low stimulus complexity. Furthermore, past studies did not investigate specific components of recognition memory such as levels of confidence or recollection and familiarity. In addition, they were guided by the notion that reinstatement should be sequential (i.e., the scan-path hypothesis), rather than the view that recognition performance might be related to content reinstatement regardless of eye fixation order.

The results of this study indicated selective effects of eye movements on the nature of recollective experience in that remember responses were associated with a higher degree of refixations than misses (Experiments 1 and 2) and know responses (Experiment 2). These effects were observed for relatively complex stimuli, while controlling for differences in number of fixations and stimulus saliency. The results extend our earlier findings involving face recognition (Mäntylä & Holm, 2006) by providing stronger and more general support for the hypothesis that the consistency of fixation distributions between study and test reflects reconstructive retrieval processes.

In Experiment 1, participants were free to move their eyes during both study and test, and hence several levels of scene information might have contributed to study-test consistency (e.g., low-level contrast as well as high-level semantic content). Despite this constraint on expected variability in terms of eye movement distribution, the refixation measure was sensitive enough to differentiate between participants' recognition judgments (as indicated by remember responses) and false rejections.

In Experiment 2, we showed that the effect size of the refixation measure was increased when saliency was experimentally controlled for. Specifically, remember responses were associated with a higher proportion of refixations than both know responses and false rejections. Importantly, remember responses were associated with an increased probability of fixating studied regions, compared to equally salient nonstudied regions. By contrast, eye movements preceding false rejections were associated with a preference for nonstudied regions over studied regions. For instance, if recognition judgments are based on sensory reconstruction,

it seems reasonable to judge the scene as new, when sensory information sampled from the scene is mostly new. Perhaps fixations on new areas result in a false hypothesis regarding the scene, and that hypothesis is inadequate in guiding saccades toward diagnostic regions. While acquiring ever new focal information, the participant becomes convinced that the scene is new, and makes a false rejection.

One interpretation of the findings in Experiments 1 and 2 is that remember responses, know responses, and false rejections reflect the output of one source of signal strength (i.e., reinstated information). In this respect, remember and know judgments would reflect different confidence criteria on a memory strength dimension (see Wixted, 2007). On this account, know responses reflect weaker memory representations than remember responses. Two underlying processes of recognition memory might still be tenable on a signal detection theory (SDT) account of recognition memory. However, the controversy regarding dual components of recognition memory refers to the underlying processes of recognition judgments, rather than to the phenomenal experience of the participants. Therefore, even if a single process SDT account of recognition memory appropriately describes current data, the application of the remember/know paradigm is motivated to the extent that these responses reflect qualitatively different recognition experiences. In addition, the SDT models of recognition memory do not predict transitions in the phenomenal quality of recognition judgments.

In fact, using eye movements might be one way of disentangling the issue of underlying processes in remember and know responses. If these responses involve the same processes and could be equated with regard to confidence (probably high confidence judgments), they should not differ in terms of eye movement consistency. Instead, if remember and know judgments reflect different underlying processes, one might expect remember responses to have a higher perceptual consistency than know responses, even when relative confidence ratings are similar.

In Experiment 3, we investigated the underlying mechanisms of reconstruction in episodic memory retrieval. Specifically, we found that episodic memory influences eye movements in scene recognition. Providing a retrieval cue (the village name) before scene presentation at test induced eye movements toward regions consistent with those fixated during scene study. This result suggests that the participants employed a hypothesis-testing process based on their expectations. This expectation expressed episodic memory of the previously encoded scene, and seems to have guided the participants' eyes toward diagnostic scene regions.

The overall findings of this study (that recognition memory is related to perceptual reinstatement) are in agreement with Loftus's (1976) account that scene memory is represented by gist and a limited number of informative scene details. Furthermore, if episodic memory influences saccade targeting, one might expect this influence to be somewhat delayed with respect to the ongoing perceptual processing (see Ballard et al., 1997). In that case, a typical case for memory driven saccades should be refixations of a specific region, followed by saccades to other regions, and then a new refixation at the original

region. The last fixation would then be a confirmatory fixation for the recognition match caused by fixating that region the first time during test. Limiting retrieval to few informative scene attributes rather than distributing fixations to several regions, or representing scenes in terms of eye movement sequences, is not only a more economic strategy, but also makes sense in light of distinctiveness effects in episodic memory.

Hunt and McDaniel (1993) defined distinctiveness as an attribute that distinguishes the item from other instances of the same general theme. Specifically, distinctiveness is thought to be the result of coding differences in the context of similarity and similarities in the context of differences (Hunt & Seta, 1984). Considering that the recognition tasks of the experiments presented here required a high degree of target-distractor discrimination, an efficient encoding and retrieval strategy according to Hunt and McDaniel's account would be to process outstanding scene features.

This notion is also consistent with Rajaram's distinctiveness/fluency account of remember/know dissociations (Rajaram, 1996; see also Mäntylä, 1997; Mäntylä & Holm, 2006). According to Rajaram, "an analysis of the distinctive or salient attributes of the information, be they conceptual or perceptual in nature, creates memories that are later accompanied by the subjective experience termed as Remember" (Rajaram, 1996, pp. 374).

The idea that long-term memory for context facilitates visual search has been shown within implicit memory (Chun & Jiang, 2003) and supports the view that long term memories can influence attention deployment in scene perception. Furthermore, recent studies by Brockmole and Henderson (2006) and Summerfield, Lepsien, Gitelman, Mesulam, and Nobre (2006) suggest that explicit memory influences attention deployment. Both studies showed that long-term memory for object locations in complex scenes facilitated object detection. To the extent that recognition of constituent parts is an important aspect of complex scene recognition, one might speculate that those same mechanisms would be involved in scene recognition by biasing attention deployment according to memory traces. Indeed, the findings by Summerfield et al. indicate that such memory influence on attention deployment in visual search can be initiated with as little as 100 msec of scene cue information. In addition, Altmann (2004) showed that people produce anticipatory saccades to regions reflecting the content of a spoken sentence before the content noun is spoken, even when the anticipated object of that sentence is no longer present at the location.

Correspondingly, our findings suggest that memory guidance is involved early, considering that the refixation measure differentiated recognition judgments within the first three fixations in all three experiments. The notion that memory govern behavior in an anticipatory way is consistent with findings in rather diverse fields of human activities, such as adjustment of grip force prior to target contact in dexterous manipulation (Gordon, Westling, Cole, & Johansson, 1993), top-down bias by word frequency in spoken word recognition (Dahan, Magnuson, & Tanenhaus, 2001; see also McClelland, 1991; McClelland & Elman, 1986), and semantic associations in visual

search (Moore, Laiti, & Chelazzi, 2003). In addition, several brain imaging studies suggest that encoding processes and their related brain activity affect which processes are reactivated during subsequent retrieval (Buckner, Wheeler, & Sheridan, 2001; Nyberg, Habib, McIntosh, & Tulving, 2000; Nyberg et al., 2001).

The present findings are also in agreement with studies showing spontaneous saccades during recall to blank regions where the retrieved information used to be (Altmann, 2004; Brandt & Stark, 1997; Laeng & Teodorescu, 2002; Spivey & Geng, 2001). Collectively, these findings suggest that episodic memory retrieval is an active, reconstructive and embodied process, where the external environment serves as an important support, or even as an "external" memory (O'Regan, 1992). Furthermore, guiding eye movements by cuing affect higher cognitive processes such as interpretation of ambiguous figures (Pomplun, Ritter, & Velichovsky, 1996), reaction time to comprehension of speech (Richardson & Dale, 2005), and problem solving performance (Grant & Spivey, 2003). This suggests that cuing attention in a scene recognition test to previously studied regions should facilitate recognition performance, whereas cuing attention to nonstudied regions should increase the risk for false rejections. Similarly, one would expect that the stronger the cue, the higher the impact on gaze control in a recognition task. For instance, the cuing effect of Experiment 3 might be expected to increase as a function of study blocks. Not only would the valid cues be more valid, but the invalid cues more invalid.

In conclusion, the present findings support the view that anticipation or hypothesis testing is an intrinsic part of retrieval in episodic recognition. It is possible that hypothesis testing is a general principle of recognition memory, characterized by an iterative testing of sensory evidence against stored representations. When that comparison is insufficient for a decision, additional information is sampled according to hypotheses regarding the sensory impression, until a judgment can be made. This active view on recognition also makes sense from a more ecological point of view. We often move about in familiar environments. Our rich experience should then provide us with the information necessary to anticipate what structures will meet us as we turn around the next corner, and this anticipation should be expressed as anticipatory deployment of attention (including gaze direction) toward upcoming regions of interest.

AUTHOR NOTE

This research was supported by a grant from the Swedish Research Council and a scholarship from the Swedish Foundation for International Cooperation in Research and Higher Education. Portions of the study were presented at the 13th European Conference on Eye Movements (ECM), Bern, 2005, and the 47th Annual Meeting of the Psychonomic Society, Houston, 2006. We are thankful for the valuable comments by Lars Nyberg in the preparation of the manuscript. Correspondence concerning this article should be addressed to L. Holm, Department of Psychology, University of Umeå, 901 87 Umeå, Sweden (e-mail: linus.holm@psy.umu.se).

REFERENCES

- ALTMANN, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: The "blank screen paradigm." *Cognition*, *93*, 79-87.

- BALLARD, D., HAYHOE, M. M., POOK, P. K., & RAO, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral & Brain Sciences*, *20*, 723-767.
- BAR, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*, 617-629.
- BAR, M., KASSAM, K. S., GHUMAN, A. S., BOSCHYAN, J., SCHMID, A. M., DALE, A. M., ET AL. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, *103*, 449-454.
- BLAXTON, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 657-668.
- BRAINERD, C. J., REYNA, V. F., & MOJARDIN, A. H. (1999). Conjoint recognition. *Psychological Review*, *106*, 160-179.
- 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*, 27-38.
- BROCKMOLE, J. R., CASTELHANO, M. S., & HENDERSON, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 699-706.
- BROCKMOLE, J. R., & HENDERSON, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition*, *13*, 99-108.
- BUCKNER, R. L., WHEELER, M. E., & SHERIDAN, M. A. (2001). Encoding processes during retrieval tasks. *Journal of Cognitive Neuroscience*, *13*, 406-415.
- CHUN, M. M., & JIANG, Y. (2003). Implicit long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 224-234.
- DAHAN, D., MAGNUSON, J. S., & TANENHAUS, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*, 317-367.
- DAVENPORT, J. L., & POTTER, M. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*, 559-564.
- DECO, G., & SCHÜRMAN, B. (2000a). A hierarchical neural system with attentional top-down enhancement of the spatial resolution for object recognition. *Vision Research*, *40*, 2845-2859.
- DECO, G., & SCHÜRMAN, B. (2000b). A neuro-cognitive visual system for object recognition based on testing of interactive attentional top-down hypotheses. *Perception*, *29*, 1249-1264.
- DEUBEL, H., & SCHNEIDER, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827-1837.
- FISHER, D. F., KARSH, R., BREITENBACH, F., & BARNETTE, B. D. (1983). Eye movements and picture recognition: Contribution or embellishment. In R. Groner, C. Menz, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and psychological functions: International views* (pp. 193-210). Hillsdale, NJ: Erlbaum.
- GARDINER, J. M., & JAVA, R. I. (1993). Recognizing and remembering. In A. F. Collins, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 163-188). Hillsdale, NJ: Erlbaum.
- GODDEN, D. R., & BADDELEY, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology*, *66*, 325-331.
- GORDON, A. M., WESTLING, G., COLE, K. J., & JOHANSSON, R. S. (1993). Memory representations underlying motor commands used during manipulation of common and novel objects. *Journal of Neurophysiology*, *69*, 1789-1976.
- GRANT, E. R., & SPIVEY, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*, 462-466.
- HEBB, D. O. (1949). *The organization of behavior*. New York: Wiley.
- HEBB, D. O. (1968). Concerning imagery. *Psychological Review*, *75*, 466-477.
- HENDERSON, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498-504.
- HOLLINGWORTH, A. (2004). Constructing visual representations from natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception & Performance*, *30*, 519-537.
- HOLLINGWORTH, A. (2006). Scene position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 58-69.

- HOLLINGWORTH, A., & HENDERSON, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 113-136.
- HOLM, L. (2007). *Predictive eyes precede retrieval: Visual recognition as hypothesis testing*. Unpublished doctoral dissertation, Umeå University.
- HUNT, R. R., & MCDANIEL, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory & Language*, **32**, 421-445.
- HUNT, R. R., & SETA, C. E. (1984). Category size effects in recall: The roles of individual item and relational information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 282-290.
- ISHIHARA, S. (1984). *The series of plates designed as a test for colour blindness*. Tokyo: Kanehara.
- JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.
- JACOBY, L. L., & CRAIK, F. I. M. (1979). Effects of elaboration at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 1-23). Hillsdale, NJ: Erlbaum.
- JACOBY, L. L., & DALLAS, M. (1981). On the relation between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.
- JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.
- KOLERS, P. A. (1973). Remembering operations. *Memory & Cognition*, **1**, 347-355.
- LAENG, B., & TEODORESCU, D.-S. (2002). Eye scanpaths during visual imagery re-enact those of perception of the same visual scene. *Cognitive Science*, **26**, 207-231.
- LOCHER, P. J., & NODINE, C. F. (1974). The role of scanpaths in the recognition of random shapes. *Perception & Psychophysics*, **15**, 308-314.
- LOFTUS, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, **3**, 525-551.
- LOFTUS, G. R. (1976). A framework for a theory of picture recognition. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes* (pp. 499-513). Hillsdale, NJ: Erlbaum.
- LOFTUS, G. R. (1981). Tachistoscopic simulations of eye fixations on pictures. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 369-376.
- MANDLER, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, **87**, 252-271.
- MANNAN, S. K., RUDDOCK, K. H., & WOODING, D. S. (1997). Fixation patterns made during brief examination of two-dimensional images. *Perception*, **26**, 1059-1072.
- MÄNTYLÄ, T. (1997). Recollection of faces: Remembering differences and knowing similarities. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1203-1216.
- MÄNTYLÄ, T., & HOLM, L. (2006). Gaze control and recollective experience in face recognition. *Visual Cognition*, **14**, 365-386.
- MELCHER, D., & KOWLER, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research*, **41**, 3597-3611.
- MCCLELLAND, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology*, **23**, 1-44.
- MCCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.
- MOORES, E., LAITI, L., & CHELAZZI, L. (2003). Associative knowledge controls deployment of visual attention. *Nature Neuroscience*, **6**, 182-189.
- MORRIS, C. D., BRANSFORD, J. D., & FRANKS, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 519-533.
- NOTON, D., & STARK, L. W. (1971). Scanpaths in eye movements during pattern perception. *Science*, **171**, 308-311.
- NYBERG, L., HABIB, R., MCINTOSH, A. R., & TULVING, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences*, **97**, 11120-11124.
- NYBERG, L., PETERSSON, K. M., NILSSON, L.-G., SANDBLOM, J., ÅBERG, C., & INGVAR, M. (2001). Reactivation of motor brain areas during explicit memory for actions. *NeuroImage*, **14**, 521-528.
- O'REGAN, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, **46**, 461-488.
- PARKER, R. E. (1978). Picture processing during recognition. *Journal of Experimental Psychology: General*, **4**, 284-292.
- POMPLUN, M., RITTER, H., & VELICHOVSKY, B. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception*, **25**, 941-948.
- RAJARAM, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 365-377.
- RAJARAM, S., & ROEDIGER, H. L., III (1996). Remembering and knowing as states of consciousness during retrieval. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to the question of consciousness* (pp. 213-240). Mahwah, NJ: Erlbaum.
- RICHARDSON, D. C., & DALE, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, **29**, 1045-1060.
- ROEDIGER, H. L., III, WELDON, M. S., & CHALLIS, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.
- SMITH, S. M. (1988). Environmental context-dependent memory. In G. M. Davies & D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 13-33). New York: Wiley.
- SPIVEY, M. J., & GENG, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, **65**, 235-241.
- SUMMERFIELD, J. J., LEPSIEN, J., GITELMAN, D. R., MESULAM, M. M., & NOBRE, A. C. (2006). Orienting attention based on long-term memory experience. *Neuron*, **49**, 905-916.
- TULVING, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
- TULVING, E. (1985). Memory and consciousness. *Canadian Journal of Psychology*, **26**, 1-12.
- TULVING, E., & THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352-373.
- WALKER-SMITH, G. J., GALE, A. G., & FINDLAY, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, **6**, 313-326.
- WIXTED, J. T. (2007). Dual-process theory and signal detection theory of recognition memory. *Psychological Review*, **114**, 152-176.
- YONELINAS, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517.

NOTES

1. We also analyzed the data with alternative measures of consistency. Specifically, we calculated the proportion of study regions being revisited during test, and a measure for sequential reinstatement based on the string edit method (see Brandt & Stark, 1997, for a complete description). As none of these measures were systematically related to recognition judgments, and they are less well motivated theoretically, they are not reported here (but see Holm, 2007).

2. Considering that looking at the same computer screen on two different occasions will produce refixations, irrespective of whether the screen content remains the same across the two occasions. Therefore, a random baseline was calculated for the refixation measure. We approximated a measure of incidental refixations by comparing eye movement data from all dissimilar study and test picture trials for each participant—that is, $36 \times (36 - 1)$ (test trials) comparisons per participant.

3. Including fast responses produced the same pattern of results as the limited set. In fact, the effect size was greater: Average values of proportions of refixations were .71 and .58 for matched and mismatched conditions, respectively, in comparison with Block 3. The difference was statistically reliable [$t(23) = 4.08, p < .001$].

(Manuscript received September 28, 2006;
revision accepted for publication January 14, 2007.)