Taking a new look at looking at nothing

Fernanda Ferreira, Jens Apel and John M. Henderson

School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh, EH8 9JZ, UK

A crucial question in cognitive science is how linguistic and visual information are integrated. Previous research has shown that eye movements to objects in the visual environment are locked to linguistic input. More surprisingly, listeners fixate on now-empty regions that had previously been occupied by relevant objects. This 'looking at nothing' phenomenon has been linked to the claim that the visual system constructs sparse representations of the external world and relies on saccades and fixations to extract information in a just-in-time manner. Our model provides a different explanation: based on recent work in visual cognition and memory, it assumes that the visual system creates and stores detailed internal memory representations, and that looking at nothing facilitates retrieval of those representations.

Looking at nothing

Normally, we look at things in the surrounding environment [1–4], but recent work using the eye movement monitoring technique has demonstrated that people will also sometimes look at nothing: that is, they will fixate on a blank location if a relevant visual stimulus had previously occupied that region of space [5–8]. Our shorthand phrase for this phenomenon will be 'looking at nothing'. Looking at nothing is of considerable interest to researchers in psycholinguistics, visual cognition and eye movement control because it lies at the interface of these different areas of cognitive science. Moreover, looking at nothing is likely to reveal fundamental aspects of the basic memory processes that take place during encoding and retrieval of information from multiple modalities, including speech, the environment, the behaviour of conversational partners and so on. The finding that people look at nothing is also consistent with our intuition that we tend to look at empty locations when we refer to an object that used to be present there. For example, when we mention something someone said, we might point to the chair in which the person previously was sitting.

Here, we discuss the looking at nothing phenomenon and use it to motivate a cognitive architecture for the integration of visual, spatial, linguistic and conceptual information. One important feature of our account is that it challenges the idea that internal representations of the visual world are sparse, amounting to little more than spatial indexes for directing eye movements to objects in the external world. This view has been invoked in past explanations of looking at nothing [5,6,8]. However, recent work in visual cognition indicates that the memory system actually creates quite detailed internal memory representations of the external world [9–11], so that information about objects, their specific visual properties and their locations are bound in a coherent memory representation. In addition, because of the tight relationship between encoding and retrieval [12–17], activation of any feature of this internal representation facilitates retrieval of the other features. We conclude that the reason people look at nothing is that re-activation of memory representations drives the eyes to previously viewed locations, and those looks to blank regions enhance subsequent memory retrieval.

Evidence that we look at nothing

There are now several independent demonstrations of looking at nothing, but we focus mainly on the two earliest reports because they have had an important influence on the way the phenomenon is currently viewed. Richardson and Spivey [6] arranged four unique faces in a 2x2 grid on a computer monitor, and each face articulated a different factual statement. After the fourth statement, the faces disappeared. Participants then heard a statement referring to one of the four facts, and their task was to say whether the statement was true. Eye movements were monitored throughout the experiment. Richardson and Spivey [6] found that people tended to look at the empty cell previously occupied by the 'fact-teller' when verifying that particular fact. To rule out the possibility that participants simply linked faces and facts, Richardson and Spivey [6] conducted a follow-up experiment using identical spinning crosses rather than faces. These stimuli yielded essentially the same pattern of results, indicating that the fact and the spatial information associated with the location from which the fact emanated were linked. However, they also observed that participants' accuracy on the fact-verification task was no higher when participants looked at the relevant cell during fact recall, a finding that led them to argue against retrieval-based accounts of looking at nothing. Richardson and Spivey [6] argued that if looking at nothing were a type of contextual memory effect, accuracy would have been higher when people looked at the previously occupied location.

A similar lack of memory facilitation was found in a related and more recent study [8]. Participants watched while a computer-animated animal apparently burrowed underground and then emerged to state a fact. The animal then either burrowed and reappeared in a way that implied

Corresponding author: Ferreira, F. (fernanda.edi@gmail.com).

1364-6613/$ – see front matter © 2008 Elsevier Ltd. All rights reserved. doi:10.1016/j.tics.2008.07.007 Available online 18 September 2008

it was the same animal (continuity condition) or its burrowing behaviour implied the appearance of a second but physically identical animal (two-animal condition). In the continuity condition, participants tended to fixate on both the original and new position of the animal; but in the two-animal condition, participants ignored the second position (i.e. the second animal), presumably because it was unconnected to the spoken fact. In both cases, the animal was no longer physically present, so the participants in the continuity condition were actually looking at a blank screen. The authors [8] argued that these results provide support for the use of an object-based attentional mechanism that enables the external world to be consulted when information must be retrieved. In addition, like Richardson and Spivey [6], the authors [8] found no link between memory accuracy and looks to the relevant blank locations.

Altmann [5] also studied looking at nothing, but he used a rather different paradigm. His previous work with Kamide and colleagues [18,19] had shown that participants look at objects as they are mentioned in spoken sentences. To investigate whether people look at blank regions previously occupied by relevant objects, participants first saw an arrangement of four objects (e.g. a man, a woman, a cake and a newspaper) for five seconds, and then the entire display went blank. After one second, a spoken sentence was presented (i.e. 'the man will eat the cake'), and the trial ended five seconds later. Eye movements were monitored throughout. Altmann [5] observed that as the subject of the sentence was spoken, participants looked at the location in the blank display previously occupied by that subject (e.g. the place where the man had been). Participants simply listened to the sentences without performing any extraneous task, and memory for sentence content was not assessed. Thus, this study does not tell us about the relationship between looking at nothing and accuracy of memory retrieval. What it does clearly demonstrate is a tight link between speech and eye movements to objects no longer present but related to the content of the utterances (see Ref. [20]).

Other investigators have reported related results. For example, during reading, readers answering questions about previously read text will tend to move their eyes back to the location in the text containing the answer, even when the text is no longer visible [21] (but see Ref. [22]). Johansson *et al.* [7] asked participants to view scenes while listening to a verbal description. When participants later described the scene from memory, their eye movements were found to be similar to the pattern executed during encoding of the scene. This tendency persisted even in complete darkness; that is, in an environment in which there was no real visual information at all, so participants were looking at nothing. Laeng and Teodorescu [23] found that scan patterns made by participants when they imagined a stimulus they had previously viewed essentially under 'blank screen' conditions were similar to those executed during encoding of the stimulus. They also observed that similarity in eye movement patterns from encoding to imagery predicted memory accuracy (contrary to Refs [6,8]), thus indicating that the eye movements made during the imagery task, which is basically a blank screen, were functional and not just a by-product of performing the task.

Why look at nothing? The world as its own memory account

Altmann [5] proposed that looking at nothing is a consequence of the activation of a spatial index associated with the mentioned object, with a saccade automatically launched to the location specified by the index [5,20]. He also suggested that the tendency to make eye movements to indexed locations arises because the visual system uses the world as an external memory rather than relying on internal representations [24,25]. Similarly to Altmann [5], Richardson and Spivey [6] proposed that viewers rely on the external world to serve as its own memory, rather than building and using an internal representation of the content of the scene. In their view [5,6], the lack of internal representation also explains why the oculomotor system would program a saccade to nothing: because there is no internal representation of what had been present before, the visual system does not detect that the display has changed from containing objects to being empty. Thus, people look at an empty region simply because the visual system directs eye movements to the external world to get information about it, rather than relying on an internal, stored representation. However, this 'world as its own memory' view is inconsistent with recent experimental evidence in visual cognition and visual memory. Before discussing this evidence, we briefly outline a contrasting theory.

Why look at nothing? A representational theory of the integration of vision and language

We propose that looking at nothing reflects something important about the nature of mental representations. Specifically, it reflects the existence of an integrated memory representation derived from visual and linguistic input [26,27]. Later, when part of an integrated representation is reactivated, the other parts are retrieved as well. This, in turn, causes the eyes to move to the location in which the item originally appeared. Moreover, returning the eyes to the former location of an object improves memory for all the information associated with that object [9,10,23], including any concepts or propositions conveyed linguistically. In this view, looking is both a consequence of the integrated representation, and serves to facilitate retrieval of further information from it. These ideas are represented in Figure 1.

To see how this might work, imagine that two people are embedded in a visual world such as the one shown in Figure 1c, and one of them says 'We hailed a taxi yesterday but no one would stop'. The utterance and the visual world together contribute to the formation of an integrated representation. Some aspects of the integrated representation are formed directly from the input (indicated with red lines). For example, the utterance leads to the direct formation of linguistic representations (including lexical meaning, syntax and prosody) in addition to conceptual representations (e.g. the word 'taxi' activates the corresponding concept). The visual world leads directly to the activation of conceptual representations too, but not linguistic representations. Unlike the utterance, the visual world generates visual representations (e.g. object shapes, colour, size and orientation) and leads to the creation of spatial location indexes [28–32] which can later be used to direct eye movements to relevant objects [11].

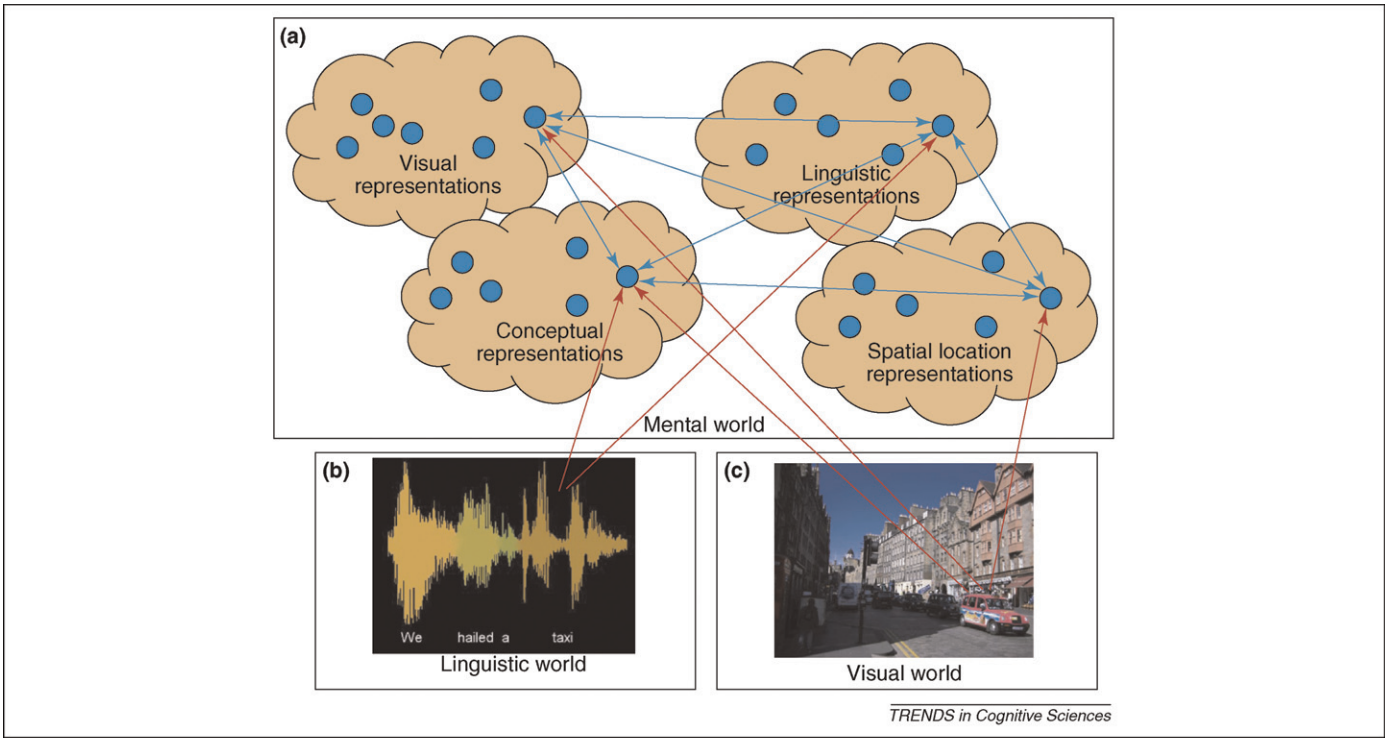

Figure 1. Model of the integration of visual and linguistic information in memory. Part **(a)** represents the mental world. The clouds are for different types of representations the cognitive system can form. Visual representations are formed from visual features; linguistic representations are formed from words and utterances; conceptual representations are pre-stored in memory, and correspond to our knowledge of concepts and categories; and spatial location representations are formed from visual input, and indicate the locations of objects in the visual world. Part **(b)** represents a spoken utterance – in this case, 'We hailed a taxi . . ..'. Part **(c)** shows the visual environment in which the interlocutors were embedded when the utterance was produced.

The blue lines in Figure 1 indicate inter-connections among internal representations. For example, an episodically related conceptual representation is associated with the corresponding visual representation (e.g. the concept 'taxicab' is associated with information about what the relevant taxi looked like). In this way, the word 'taxi' might later activate the corresponding stored visual representation of the taxi that was viewed in the same context. Figure 1 depicts all internal representations as interconnected. Thus, the experience of hearing the sentence 'We hailed a taxi...' and seeing the street scene, leads to an integrated representation containing visual, conceptual, linguistic and spatial location information. Retrieval of any piece of this representation facilitates retrieval of the other pieces [12,13,17]. If the word 'taxi' were later presented, it would lead to the retrieval of information about the sentence, the visual properties of the image, the concepts associated with both and the locations of the objects. Our theory also predicts that simply moving the eyes to the location occupied previously by, say, the taxi will activate the other information as well. As a result, looking at the empty location will facilitate memory for both the scene and the utterance content.

Our proposal of an integrated multi-modal memory representation is related to the idea of event files [33], which is in turn related to the original idea of object files [34] (see also Refs [35–37]). Object files are integrated perceptual and conceptual representations that contain an episodic trace indexed by spatio-temporal information. There is substantial evidence that retrieval from memory is facilitated when spatio-temporal continuity or relative location is maintained [33–37]. We extend this notion with the proposal that during spoken language comprehension, people create a bound and integrated representation from visual and linguistic input, and that because of principles such as encoding specificity and possibly also re-enactment, retrieval for one part of this representation is facilitated when another part is reinstated [12,38]. Thus, eye movements to the appropriate region of the blank screen are made because an integrated representation was formed during encoding, and when it is time to retrieve information about some part of the stimulus, retrieval is made easier if spatial information matches. The eyes move to where the object was both as a cause and as a consequence of memory retrieval.

Evidence for integrated representations: recent work in visual cognition

The integrated representation theory assumes that people form detailed internal representations of the external world. However, this idea runs counter to the view that 'the world is its own memory' [24,25,39,40]. The earliest demonstrations of 'change blindness' were taken as evidence for this notion. These studies show that people can be insensitive to large visual changes to objects right in front of their eyes [41–43]. Additional support for the idea that the visual system relies only on spatial pointers to external objects but does not represent the objects themselves [39], came from research using a block copying task [44]. People tended to look at one block to encode a relevant feature such as colour, then look again to encode location, and so on, indicating that they did not form a memory representation containing information about multiple features (see also Ref. [45]).

However, the notion that the cognitive system does not construct internal representations of the visual world has recently been undermined [10,11,46,47]. It is now clear that people do, in fact, encode and retain a great deal of information about the objects on which they fixate [9–11]. Even more compelling are studies showing that participants retain detailed information in their long-term memory about objects that they are not intentionally trying to encode or remember [48,49]. People do sometimes adopt strategies in block-copying [45,50] or scene-comparison tasks [45] that seem to involve encoding just one object or even one visual feature at a time, but this is because of strategic decisions to minimize working memory load rather than an inability of the cognitive architecture to retain an internal representation [51,11].

Finally, we believe there is a serious problem with the 'world as its own memory' account of looking at nothing when it involves spoken language. The problem is that the world cannot serve as its own memory for 'auditory' information. That is, the 'visual' world can be re-sampled with fixations because objects in the environment tend to stay put. However, if you hear an utterance and then move your eyes to the location from which it emanated, you will not succeed in recreating the stimulus. The eye movements, then, must be useful for reactivating an 'internal' representation of the utterance, as our model assumes. In our view, looking to a location will facilitate retrieval of auditory information from memory, and, therefore, looking at nothing might facilitate memory for spoken linguistic content.

As mentioned at the beginning of this article, several studies show a relationship between eye movements and memory performance. For example, Laeng and Teodorescu [23] found that similarity in scan paths from an encoding condition to a blank-screen imagery condition predicted memory accuracy from the encoding condition. Several studies in the change blindness literature have shown that change detection is facilitated when the eyes spontaneously return to the changed region compared to when they do not [9,42]. This latter effect has also been found when attention is directed to the changed region via an attentional cue [10,52]. The previously mentioned object and event file literatures show facilitated memory retrieval when the retrieved information comes from the same location or the same object if its location is updated via motion [33–37]. Thus, there is considerable evidence in the literature that attending to and looking at a location facilitates retrieval of information which had occupied that location.

The theory summarized in Figure 1 assumes that when memory is probed, spatial indexes are retrieved and the activation of those spatial indexes increases the probability of an eye movement to the corresponding spatial location. Recent studies show that memory for object information is enhanced when an object is probed in its original spatial location. Hollingworth [52] compared memory for an object previously viewed within a scene when the object was tested in the same or in a different scene location. He found that people were more accurate at remembering both the specific details of the object and its orientation when the object at test was presented in its original location in the scene. These results are consistent with other studies, demonstrating that attending to a region in a real-world scene facilitates retrieval of previously encoded information from that region [53,54]. It, therefore, seems that memory for objects and their properties is enhanced when spatial information at encoding, and at test, match. These findings, thus, support a fundamental assumption of our view: integrated representations including spatial indexes are naturally formed during scene processing. In a situation in which a person listens to speech in the context of a relevant scene, information from the utterance, the objects in the scene and the location of those objects are all likely to be part of such bound memory representations. Therefore, making eye movements to a now empty region that was previously occupied by a relevant object will facilitate memory performance because the location information is linked to utterance and scene content.

Given our account, why did Richardson and Spivey [6] not find better memory performance for fixated empty cells? We believe their null effect was probably because of an item selection artefact. In the Richardson and Spivey [6] experiments, trials were sorted post hoc into fixate-on-correct-square versus fixate-on-incorrect-square categories. This procedure does not control for item difficulty across the two fixation 'conditions', and so it is possible (in fact, quite likely) that facts which were more difficult to verify triggered more saccades to the location linked to the fact than did facts which were easy to verify. This would lead people to make fixations to the empty regions more often for the difficult items, and less often for the easier ones. However, it would still be true that the more difficult items are more difficult, and so accuracy might still be lower for them. Thus, the null effect associated with looking or not looking at the right empty square could be the result of two offsetting tendencies: looking at the correct empty square helps performance; but those same items are more difficult, which depresses performance. A similar explanation can account for the lack of correlation between looking behaviour and memory performance in Hoover and Richardson [8].

Looking at nothing links the internal and external worlds

The initial creation of a representation including spatial codes leads to future looks when an object is re-mentioned. Looking at nothing facilitates memory retrieval for other information associated with that location, including visual details and linguistic content. Moreover, the mechanism that leads to facilitated memory when one looks at a previously occupied location – namely, the existence of an integrated representation in which multiple features including location have been encoded – also generates the looks in the first place. In this view, when a location containing an object is initially attended to and fixated, its features are bound together in an integrated trace. When some part of that trace is later cued, the rest of the trace is retrieved including spatial information, which drives the eyes back to that location. An additional

Box 1. Outstanding questions

- What is the nature of spatial indexes, and to what extent does activating them automatically versus intentionally drive the eyes to the corresponding location?
- Are people more likely to fixate blank regions when they are trying to retrieve information that is more difficult?
- How do our theoretical framework and the looking at nothing phenomenon relate to reading behaviour, and in particular, the targeting of regressive eye movements?
- What happens to people's tendency to look at nothing when they are immersed in dynamic environments?
- Are linguistic representations in memory affected by the co-presence of visual information, and are visual representations in memory affected by the co-presence of linguistic information?
- How do visual saliency and linguistic saliency map on to each other?
- How stable are the integrated representations constructed from visual and linguistic information? How long do they last?

consequence of having created a bound representation is that once the look takes place, the location then serves as an additional retrieval cue that further facilitates access of additional aspects of the original trace. At times, there might even be an intentional component to this looking behaviour: people could learn that attending to locations facilitates memory retrieval, and then use that knowledge to strategically enhance memory by re-fixating previously occupied locations. However, whether the looks are intentional or are unconsciously triggered, the conclusion is the same: looking at nothing is an entirely expected consequence of human cognitive architecture (Box 1).

References

1 Buswell, G. (1935) *How People Look at Pictures,* University of Chicago Press
2 Yarbus, A. (1967) *Eye Movements and Vision,* Plenum Press
3 Rayner, K. (1998) Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422
4 Henderson, J.M. (2003) Human gaze control in real-world scene perception. *Trends Cogn. Sci.* 7, 498–504
5 Altmann, G.T. (2004) Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition* 93, B79–B87
6 Richardson, D.C. and Spivey, M.J. (2000) Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition* 76, 269–295
7 Johansson, R. *et al.* (2006) Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cogn. Sci.* 30, 1053–1079
8 Hoover, M.A. and Richardson, D.C. (2008) When facts go down the rabbit hole: contrasting features and objecthood as indexes to memory. *Cognition* 108, 533–542
9 Hollingworth, A. and Henderson, J.M. (2002) Accurate visual memory for previously attended objects in natural scenes. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 113–136
10 Hollingworth, A. (2008) Visual memory for natural scenes. In *Visual Memory* (Luck, S.J. and Hollingworth, A., eds), pp. 123–162, Oxford University Press
11 Henderson, J.M. (2008) Eye movements and visual memory. In *Visual Memory* (Luck, S.J. and Hollingworth, A., eds), pp. 87–121, Oxford University Press
12 Thompson, D.M. and Tulving, E. (1970) Associative encoding and retrieval: weak and strong cues. *J. Exp. Psychol.* 86, 255–262
13 Tulving, E. and Thompson, D.M. (1973) Encoding Specificity and retrieval processes in episodic memory. *Psychol. Rev.* 80, 352–373
14 Roediger, H.L. and Guynn, M.J. (1996) Retrieval processes. In *Handbook of Perception and Cognition: Memory* (Bjork, E.L. and Bjork, R.A., eds), pp. 197–236, Academic Press
15 Godden, D.R. and Baddeley, A.D. (1975) Context-dependent memory in two natural environments: on land and underwater. *Br. J. Psychol.* 66, 325–331
16 Eich, J.E. (1980) The cue-dependent nature of state-dependent retrieval. *Mem. Cognit.* 8, 157–173
17 Kent, C. and Lamberts, K. (2008) The encoding-retrieval relationship: retrieval as mental simulation. *Trends Cogn. Sci.* 12, 92–98
18 Altmann, G.T. and Kamide, Y. (1999) Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73, 247–264
19 Kamide, Y. *et al.* (2003) The time-course of prediction in incremental sentence processing: evidence from anticipatory eye-movements. *J. Mem. Lang.* 49, 133–156
20 Altmann, G.T. and Kamide, Y. (2007) The real-time mediation of visual attention by language and world knowledge: linking anticipatory (and other) eye movements to linguistic processing. *J. Mem. Lang.* 57, 502–518
21 Kennedy, A. and Murray, W.S. (1987) Spatial coordinates and reading: comments on monk (1985). *Q. J. Exp. Psychol.* 39A, 649–656
22 Inhoff, A.W. and Weger, U.W. (2005) Memory for word location during reading: eye movements to previously read words are spatially selective but not precise. *Mem. Cognit.* 33, 447–461
23 Laeng, B. and Teodorescu, D.S. (2002) Eye scanpaths during visual imagery re-enact those of perception of the same visual scene. *Cogn. Sci.* 26, 207–231
24 O'Regan, J.K. (1992) Solving the 'real' mysteries of visual perception: the world as an outside memory. *Can. J. Psychol.* 46, 461–488
25 O'Regan, J.K. and Nöe, A. (2001) A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–1011
26 Jackendoff, R. (1996) The architecture of the linguistic-spatial interface. In *Language and Space* (Bloom, P. *et al.*, eds), pp. 1–30, MIT Press
27 Knoeferle, P. and Crocker, M.W. (2007) The influence of recent scene events on spoken comprehension: evidence from eye movements. *J. Mem. Lang.* 57, 519–543
28 Horowitz, T.S. and Thornton, I.M. (2008) Objects or locations in vision for action? Evidence from the MILO task. *Vis. Cogn.* 16, 486–513
29 Pylyshyn, Z. (2004) Some puzzling findings in multiple object tracking (MOT): I. Tracking without keeping track of object identities. *Vis. Cogn.* 11, 801–822
30 Pylyshyn, Z. (2000) Situating vision in the world. *Trends Cogn. Sci.* 4, 197–207
31 Scholl, B.J. (2001) Tracking multiple objects through occlusion: clues to visual objecthood. *Cognit. Psychol.* 38, 259–290
32 Zelinsky, G.J. and Neider, M.B. (2008) An eye movement analysis of multiple object tracking in a realistic environment. *Vis. Cogn.* 16, 553–566
33 Hommel, B. (2004) Event files: feature binding in and across perception and action. *Trends Cogn. Sci.* 8, 494–500
34 Kahneman, D. *et al.* (1992) The reviewing of object files: object-specific integration of information. *Cognit. Psychol.* 24, 175–219
35 Henderson, J.M. (1994) Two representational systems in dynamic visual identification. *J. Exp. Psychol. Gen.* 123, 410–426
36 Gordon, R.D. and Irwin, D.E. (1996) What's in an object file? Evidence from priming studies. *Percept. Psychophys.* 58, 1260–1277
37 Mitroff, S.R. *et al.* Staying in bounds: contextual constraints on object file coherence. *Vis. Cogn.* (in press)
38 Kolers, P.A. (1976) Pattern-analyzing memory. *Science* 191, 1280–1281
39 Ballard, D.H. *et al.* (1997) Deictic codes for the embodiment of cognition. *Behav. Brain Sci.* 20, 723–767
40 Rensink, R. (2000) The dynamic representation of scenes. *Vis. Cogn.* 7, 17–42
41 Grimes, J. (1996) On the failure to detect changes in scenes across saccades. In *Perception: Vancouver Studies in Cognitive Science* (Vol. 5) (Akins, K., ed.), pp. 89–110, Oxford University Press
42 Henderson, J.M. and Hollingworth, A. (1999) The role of fixation position in detecting scene changes across saccades. *Psychol. Sci.* 10, 438–443
43 Rensink, R.A. *et al.* (1997) To see or not to see: the need for attention to perceive changes in scenes. *Psychol. Sci.* 8, 368–373
44 Ballard, D.H. *et al.* (1995) Memory representations in natural tasks. *J. Cogn. Neurosci.* 7, 66–80

45 Gajewski, D.A. and Henderson, J.M. (2005) Minimal use of working memory in a scene comparison task. *Vis. Cogn.* 12, 979–1002
46 Henderson, J.M. and Castelhano, M.S. (2005) Eye movements and visual memory for scenes. In *Cognitive Processes in Eye Guidance* (Underwood, G., ed.), pp. 213–235, Oxford University Press
47 Clark, A. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press (in press)
Q2
48 Castelhano, M.S. and Henderson, J.M. (2005) Incidental visual memory for objects in scenes. *Vis. Cogn.* 12, 1017–1040
49 Williams, C.C. *et al.* (2005) Incidental visual memory for targets and distractors in visual search. *Percept. Psychophys.* 67, 816–827
50 Hayhoe, M. *et al.* (1998) Task constraints in visual working memory. *Vision Res.* 38, 125–137
51 Hayhoe, M. and Ballard, D. (2005) Eye movements in natural behavior. *Trends Cogn. Sci.* 9, 188–193
52 Hollingworth, A. (2006) Scene and position specificity in visual memory for objects. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 58–69
53 Hollingworth, A. (2003) Failures of retrieval and comparison constrain change detection in natural scenes. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 388–403
54 Hollingworth, A. (2005) Memory for object position in natural scenes. *Vis. Cogn.* 12, 1003–1016