

On the Functional Role of Implicit Visual Memory for the Adaptive Deployment of Attention Across Scenes

Marvin M. Chun

Yale University, New Haven, CT, USA

Ken Nakayama

Harvard University, Cambridge, MA, USA

Although very little visual information is explicitly retained across views, some continuity of processing is afforded by implicit visual memory traces of previous views. These memory traces interact with attentional mechanisms to guide eye movements, cognition, and action. Two different memory mechanisms are described here. First, the deployments of focal attention and eye movements are facilitated towards recently attended features and locations (priming of popout). Second, attention is guided by implicit memory traces of specific visual contexts experienced in the past (contextual cueing). Compared to the visual memory tapped by change blindness tasks, the implicit memory mechanisms of priming of popout and contextual cueing do not require conscious intervention and may exhibit greater memory capacity, longer durability, and higher discriminability. Thus, these implicit traces of past views guide attention and eye movements to allow for effective access (indexing) to a scene's details, hence providing context and continuity to ongoing interactions with the perceptual world.

Detailed visual representations of the world are clearly ephemeral (Averbach & Coriell, 1961; Neisser, 1967; Sperling, 1960) and fail to persist from one view to another, contrary to one's intuitions that our rich visual experience reflects a seamless integration of a scene's details over eye movements and over time. This failure is highlighted in work on *change blindness*, which is a striking inability to detect changes to objects and scenes from one view to the next. The pervasiveness of this impairment has been illustrated in an admirable diversity of paradigms. Thus, change blindness occurs for changes introduced to

Please address all correspondence to M.M. Chun, Department of Psychology, Vanderbilt University, 531 Wilson Hall, Nashville, TN 37240, USA. Email: marvin.chun@yale.edu

The authors thankfully acknowledge helpful feedback from Dan Simons, John Henderson, and an anonymous reviewer. This research was supported by a Yale University Social Science Faculty Development Award to MMC, and AFOSR, McKnight Foundation, and HFSP grants to KN.

stimulus arrays or scenes with an intervening interval (Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974; Simons, 1996), across eye movements (Carlson-Radvansky & Irwin, 1995; Grimes, 1996; Henderson, 1997; Irwin, 1991; McConkie & Currie, 1996), across alternating images (Rensink, O'Regan, & Clark, 1997), and even in motion picture sequences across movie cuts (Levin & Simons, 1997).

The extreme poverty of visual representations across views poses a host of puzzling challenges to the understanding of perception. If visual details are not preserved, then this means that the smooth, continuous percept of a stable world across eye movements cannot be based on integration of the sequence of images in a hypothetical visual buffer (McConkie & Rayner, 1976; Nakayama, 1990). Lack of transsaccadic integration would wreak havoc for correspondence mechanisms linking objects and events across images in a hypothetical metric-preserving representation (O'Regan & Levy-Schoen, 1983). And the failure to benefit from prior processing could possibly overburden the visual system by presenting it with a "new" scene at every single gaze.

The intriguing, alternative view is that poor visual memory for details may not be problematic at all. Instead it can be interpreted as affording parsimony to the visual system. According to this view, the lack of integration is beneficial for perceiving continuity across views as most of the irrelevant details which would otherwise overload the system are discarded (Irwin, 1996; O'Regan & Levy-Schoen, 1983; Simons & Levin, 1997). Change blindness also supports a more extreme view stating that the visual details of the world do not need to be represented in the perceiver at all (Dennett, 1991; J.J. Gibson, 1966; MacKay, 1973). O'Regan (1992) articulated this in stating that the perceived richness of the visual world is an illusion, and the detailed visual information resides in the world itself rather than in some representation within the head. Thus, the information does not need to be represented because perceivers have immediate access to visual details by simply casting their eyes (or attention) upon the external world which "serves as its own memory".

There is great merit in the idea that the brain should not represent more information than it needs to, a point elegantly demonstrated by Ballard, Hayhoe, and Pelz (1995; also see review by Hayhoe, this issue). They analysed the eye movements people make while performing a very natural block copying task that recruits visual memory for the model to be copied. Critically, subjects appear to serialize the visual memory component of the task by making repeated eye movements to the model. This minimizes the number of items that need to be stored in explicit short-term memory. Moreover, such sampling processes appear to be amnesic in a variety of other routine visual behaviors. For example, Horowitz and Wolfe (1998) recently demonstrated that search processes do not benefit from extended viewing, suggesting that attentional mechanisms do not keep careful record of locations they have visited. In sum, change blindness and these related paradigms reveal an ecological heuristic for dealing

with stimulus information overload. Internal memory is expensive so the demands on it are kept to a bare minimum, and access to a scene's content occurs as a serial sampling process of limited bandwidth (Nakayama, 1990).

Yet, there are obvious perils and limitations in a system without any memory or some internal representation. Granted that perception at the focus of attention is clear and distinct, what determines how the "outside" memory is to be sampled in the future? How can attention and eye movements be redirected to just the right part of the scene?

Remarkably, attention and eye movements move purposefully and adaptively from one region of interest to the next with few wasted fixations (see Hayhoe, this issue). Consider Fig. 1, which presents an elegant analysis of eye movement patterns obtained while observers viewed natural scenes (Yarbus, 1967). This well-known demonstration illustrates how observers focus on important details of an object or scene (see Fig. 1a), and how the set of fixated regions is sensitive to task demands (Fig. 1b & 1c). In the present discussion, however, we stress another remarkable aspect of these eye movement patterns. Namely, they are accurate and selective. The eye movements do not randomly meander about the scene, nor are there many wasted fixations. The direction of eye movements is purposeful and sophisticated. Also note the highly repetitive nature of the fixations. Although the deployment of eye movements over a scene is focused and selective, the number of times a fixation lands on an object or region of interest is extremely repetitive and seemingly redundant. This is reminiscent of the saccade pattern observed in Ballard et al.'s (1995) block copying task, and this example shows that repetitive sampling occur even in the absence of an explicit visuomotor task. Thus, these results suggest that perceivers continually interrogate the image for information and that the visual system relies on frequent access rather than internal memory to guide cognition and action.

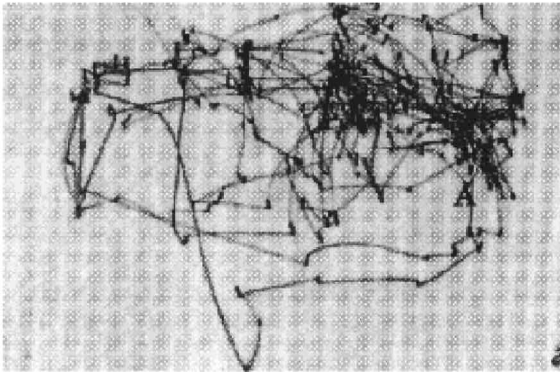
If the world serves as its own memory, effective access is imperative. The informational details in the visual world can be accessed by a set of sophisticated "deictic" primitives, which are markers or pointers to task-relevant aspects of a scene (Ballard et al., 1995; Hayhoe, this issue; Nakayama, 1990; O'Regan, 1992; Pylyshyn, 1989; Rensink, this issue). These permit internal referencing to distal features important for visual operations and action. But the fundamental issue is how does the system go about setting up these pointers?

Previous work suggests a number of possible ways to accomplish efficient indexing (see also Rensink, this issue). First, in natural scenes, a thread of context and set of deictic primitives may be provided by the gist and meaning of scenes which are quickly extracted and maintained across views (Simons & Levin, 1997). This is possible because the gist of a scene is readily available within the first few hundred milliseconds of presentation (Biederman, Mezzanotte, & Rabinowitz, 1982; Intraub, 1981; Potter, 1975). Second, in relation to gist, abstract object identity information of specific items is integrated

(a)



(b)



(c)

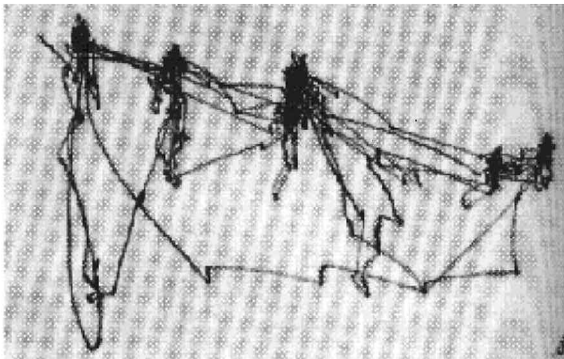


FIG. 1. Eye movement patterns (from Yarbus, 1967). (a) A picture is shown to subjects and their eye movements are recorded while they were asked to estimate the wealth of the family (b) or the ages of the people (c). The fixations are selectively deployed in a redundant manner to regions of interest within the scene.

across eye movements. For example in reading tasks, parafoveal previews of postsaccadic target words or line drawings speeded naming response times (Pollatsek, Rayner, & Collins, 1984; Rayner, McConkie, & Zola, 1980). Researchers proposed that object file representations (limited capacity visual short-term memory) mediate the perception of stability and continuity across scene changes and eye movements (Henderson, 1997; Henderson & Anes, 1994; Irwin, 1992a, b, 1996; Kahneman, Treisman, & Gibbs, 1992; Rensink, this issue). Finally, an extensive catalogue of “bottom-up” image properties serve to attract and guide attention (see Wolfe, 1998, and Yantis, 1998, for recent reviews). These do not reference memory traces of past images.

However, we argue that gist, object files, and bottom-up search mechanisms are *not* sufficient for understanding how visual processing can be planned and executed in an adaptive manner. More *specific* information (beyond gist) needs to be *retained* (beyond bottom-up processing) from one moment to the other to enable efficient interactions with the visual world. Most of the images we encounter are too complex and unique for a scene’s gist or schema to provide precise guidance to focal attention mechanisms. Object files may permit more refined guidance, but transsaccadic object-specific preview effects tend to become diluted in the presence of other objects (Henderson & Anes, 1994; Irwin, 1992a), limiting their potential utility in complex scenes. Finally, powerful, bottom-up visual routines are no doubt important, but we argue that the need to create new representations during each fixation to guide attention may impose an unnecessary burden on visual processing mechanisms. The large amount of redundancy that exists from one view to the next is likely to benefit visual processing in some fashion, even if it may not support an ability to explicitly detect changes.

Hence, our goal is to explore strategies that may benefit visual processes that retain very little conscious information across views. Our proposal is that this is accomplished by a set of independent, “implicit memory” mechanisms that allow specific visual information from scenes to persist across image changes and over time. These memory traces guide attention to ensure its efficient deployment. Such mechanisms are not necessarily under conscious control nor does the observer need to have explicit access to the underlying content of the visual representations.

We discuss two different mechanisms that permit efficient deployment of attention based on information extracted from previous views of scenes. These mechanisms are independent of each other, but work in concert to guide visual processing. First, *priming of popout* is an important, transient mechanism that facilitates the deployment of attention on a moment to moment, fixation to fixation basis (Maljkovic & Nakayama, 1994, 1996, in press). This allows image features to guide the efficient redirection of attention and eye movements without requiring the intervention of effortful and conscious decision making. Second, visual context information from past views also guides attention, and this

facilitation is termed *contextual cueing* (Chun & Jiang, 1998, 1999). Such memory traces for context accumulate over both short-term and long-term perceptual experience, and like the representations driving priming of popout, these traces need not be explicitly recognizable to guide visual behaviour.

PRIMING OF POPOUT: GUIDANCE FROM PREVIOUSLY ATTENDED FEATURES

In their studies on the *priming of popout (PoP)*, Maljkovic and Nakayama (1994, 1996, in press) have shown that the rapid deployment of attention to an odd target is facilitated by events in the very recent past. Using a display similar to that used in visual search studies but requiring focal as opposed to distributed attention (Bravo & Nakayama, 1992), they asked observers to discriminate the shape of the odd-coloured target (see Fig. 2a). Analysing the latency of correct discriminations in terms of whether the current target colour matched those appearing at various distant trials in the past, they demonstrated the existence and characteristics of a specialized memory system useful for the deployment of attention. Most important was the fact that the memory had a short time course, decaying in an approximately exponential fashion over a period of approximately 30sec (over eight trials for the customary intertrial intervals used). The data in Fig. 2b shows the difference in response latency for targets which had the same as opposed to different colours, plotted in terms of past and future trials. Note that trials in the immediate past have the greatest influence with the effects of the memory decaying continuously over time.

Maljkovic and Nakayama found that this short-term priming has many other defining characteristics: (1) It cannot be mediated by explicit conscious processing because explicit memory for trials so far back is absent (Maljkovic & Nakayama, in press). (2) It is passive and autonomous, it cannot be influenced by prior knowledge or preparation for attending to the colour of upcoming trials. (3) It can also occur for the position of a target such that targets appearing in the same position within a period approximately eight trials in the past will speed up correct responding (similar to that depicted in Fig. 2a). Note that position priming is largely landmark or object centred, so it is not tied solely to the retinal position. (4) It is restricted to the target and distractor feature that guides focal attention (Maljkovic & Nakayama, 1994). (5) It is cumulative, such that many repetitions of the same target yield even greater reductions in reaction time.

To confirm its role in the direction of attention and to link it more specifically to saccadic eye movements, McPeck, Maljkovic, and Nakayama (in press) conducted further experiments using saccadic latencies as a dependent measure. If the programming of saccadic eye movements requires attention as originally postulated by Posner (1980) and Fischer (1987) and supported by subsequent studies (Henderson, Pollatsek, & Rayner, 1989; Hoffman &

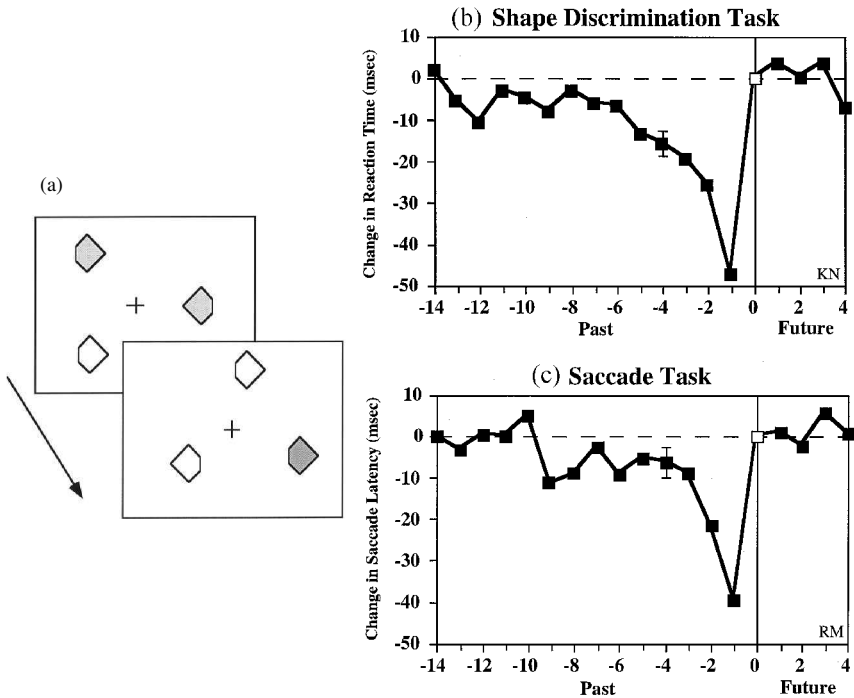


FIG. 2. Priming of popout. (a) Target arrays used to reveal short-term memory in a popout task. The task of the observer is to maintain fixation and to respond by quickly identifying the shape of the odd-coloured target as it changes colour and position in a random fashion throughout a sequence of trials. (b) For varying positions in time relative to a given trial (abscissa), there is a heavy dependence on whether such neighbouring trials have the same or different target colours. On the ordinate is plotted the difference between different vs. same trial colours. The approximately exponential function indicates that the occurrence of a same-colour trial at any time in the recent past (up to eight trials ago) can speed performance. (c). Same, except the task is for the observer to make a saccadic eye movement to odd-coloured target, and no response as to the shape of the odd-coloured target is required. A similar short-term memory function is evident.

Subramaniam, 1995; Kowler, Anderson, Doshier, & Blaser, 1995; Shepherd, Findlay, & Hockey 1986), then we might expect a similar pattern of results were we to require that the observer simply make an eye movement to the odd-coloured target, even when he or she is not required to do any additional task. McPeck et al. used an array similar to that shown in Fig. 2a and measured eye movement latencies and their dependence on target colours in the past. These results show a similar pattern of implicit memory as well, closely mirroring the behaviour observed in the discrimination experiment. Fig. 2c presents saccadic eye movement latencies, analysed in the same manner as in Fig. 2b. Note that an analogous set of results is also found. Saccadic latencies are faster to coloured targets if that trial is preceded by recent trials having the same colour.

These findings suggest that the directing of eye movements and the deployment of focal attention on which it so heavily depends, are selectively guided towards items recently attended to in the past and away from things recently avoided. As such, it provides at least some explanation as to why we can make so many eye movements (3–4 per sec) and not be thinking much about their exact destination, as suggested by Yarbus' analysis of eye movement patterns (Fig. 1). PoP facilitates the redundant, redeployment of attention and eye movements to regions of scenes that were relevant to behaviour in previous views.

During interactions with real scenes, PoP could work together with several other mechanisms that compete to guide attention and eye movements. First and foremost, it is important to remember that PoP guidance may be overridden by higher-level attentional control strategies more closely tied to the immediate goals of the perceiver. For instance, detection of popout features is subject to task-dependent attentional control settings (Bacon & Egeth, 1994; Folk, Remington, & Johnston, 1992).

Second, PoP is in direct competition with other mechanisms known to influence attentional allocation in an opposite manner. For instance, inhibition of return (IOR) is a possible mechanism that could bias the redeployment of attention away from previously attended locations (Klein, 1988; Posner & Cohen, 1984). IOR may facilitate attending to novel objects and events, opposite to the facilitation for old objects produced by PoP. The two mechanisms of PoP and IOR may competitively interact with each other, with the "winner" determined by a host of stimulus and task factors. One important parameter appears to be their time course. The influence of PoP persists across intervening visual events, whereas IOR is usually only observable on successive visual episodes. Indeed, IOR has been inferred in PoP paradigms when examining the effects of repetition at trial N-1. This was suggested by a smaller PoP effect at trial N-1 relative to trial N-2, in contrast to the monotonic time course of PoP observed in Fig. 2 (Maljkovic & Nakayama, 1996). In combination, these mechanisms could plausibly lead to the observed dynamic of the eyes darting away, then returning to recently fixated features.

The third issue we consider is whether PoP is restricted to popout targets or whether the phenomenon is more widespread. The latter view is supported by a recent study by Bichot and Schall (1999) using conjunctive visual search. In these tasks, no salient target pops out, yet a very similar priming effect can be observed for previously presented targets.

Fourth, we introduce PoP here as one of many types of intertrial facilitation effects that have recently been reported in the literature. Found and Müller (1996) demonstrated that attending to a target defined along a particular feature dimension such as orientation facilitated detection of a popout target defined by that same dimension on subsequent trials (see also Treisman, 1993). Other findings show that intertrial priming effects are not limited to popout targets. One common finding is that viewing an object facilitates subsequent

identification of that object across changes in size and viewpoint (Biederman & Cooper, 1992; Biederman & Gerhardstein, 1993; but see Tarr and Bülthoff, 1995). Using a different paradigm, DeSchepper and Treisman (1996) reported intertrial repetition and negative priming effects for novel distractor shapes that were inhibited during target selection. Ignoring a distractor on one trial made it easier to ignore and harder to attend to the same item on subsequent encounters. These findings are thus related to the PoP mechanisms described here.

CONTEXTUAL CUEING: GUIDANCE FROM INVARIANT GLOBAL VISUAL CONTEXT

In addition to the temporary valencing of loci and features, the global visual context of a scene can also guide the deployment of attention. The visual environment is extremely rich and complex, presenting constant information overload to perceivers (Broadbent, 1958). But the visual world also contains structure in the form of redundancies and regularities, as objects and events tend to covary with each other in an invariant manner over time. For instance, cars can be found on roads, windows on buildings, traffic lights at intersections, and so on. Sensitivity to such covariational information may serve to decrease complexity and increase predictability (Biederman et al., 1982; E.J. Gibson, 1969). This invariant, covariational information is present in the visual context of global scenes and images. Hence, it is plausible that sensitivity to such contextual information experienced in past views could benefit visual processing.

Chun and Jiang (1998) recently examined how contextual information guides visual attention in a paradigm called *contextual cueing*. Subjects simply performed visual search tasks of moderate difficulty, such as searching for a rotated T amongst rotated Ls. Performance is dependent on the number of items in the display, suggesting that focused attention is needed to perform the task. A role for context influencing search was examined as follows. First, arbitrary, novel visual contexts were created to avoid problems associated with defining and manipulating the “context” of natural scenes. This was achieved by operationalizing global context as the spatial layout of objects in the visual search displays. Two sample configurations are shown in Fig. 3a. The global contexts were made invariant by simply repeating sets of these configurations across blocks of trials throughout the experimental session. These are referred to as Old contexts. Finally, the contexts were predictive of target location, the critical variable in search tasks. In other words, targets appeared in consistent locations within their invariant contexts. Hence, sensitivity to the invariant configurations was predictive of the target location for that trial. The question then is, are subjects sensitive to this contextual information? If so, search would be facilitated for targets appearing in Old contexts, relative to targets appearing in New contexts.

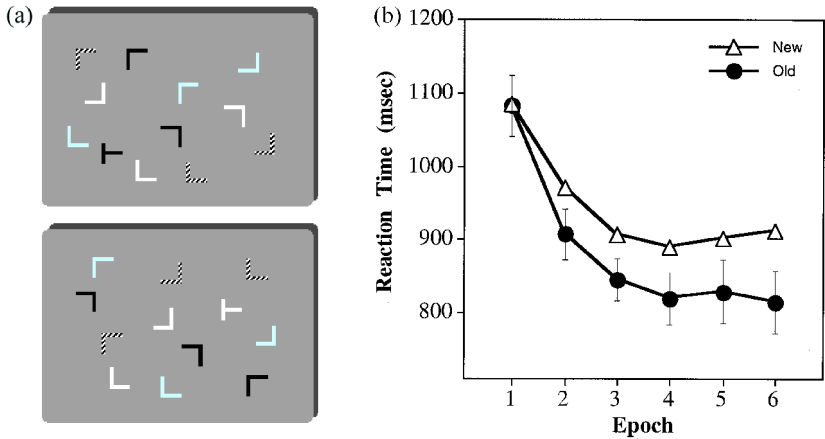


FIG. 3. Contextual cueing. (a) Two sample search arrays depicting different configurations (Old condition) which were repeated once per block throughout the entire experiment. The set of Old configurations was intermixed with randomly generated New configurations serving as baseline. The task was simply to locate and determine the orientation of the single rotated T amongst L distractors. (b) Search performance as a function of epoch (groups of five blocks) was faster for targets appearing in Old configurations vs. targets appearing in New configurations. This contextual cueing benefit emerged with experience and was significant by Epoch 2. The memory traces are implicit as subjects are at chance at discriminating Old patterns from New.

The results are shown in Fig. 3b (Chun & Jiang, 1998, exp. 1). Target discrimination RT is shown for both Old and New conditions as a function of epoch. As subjects performed the task, performance for targets in each condition diverged, revealing faster search for targets appearing in Old displays. This indicates that subjects were learning to discriminate Old contexts from New, and that this contextual information was used to guide spatial attention in the search task. This guidance was, by definition, instance based (Logan, 1988). Remarkably, this learning occurred despite the fact that subjects were never asked to try to encode the contexts, nor were they informed of the covariation between contexts and target locations. Thus, this benefit represents a form of implicit learning (Lewicki, Hill, & Czyzewska, 1992; Reber, 1989; Stadler & Frensch, 1998). Moreover, these contexts guided spatial attention despite the fact that subjects could not discriminate Old from New contexts in a forced-choice explicit recognition test. Hence, contextual knowledge was implicit (Schacter, 1987). Finally, contextual cueing can be distinguished from low-level repetition priming (Tulving & Schacter, 1990). Contextual cueing was not obtained when the target locations did not covary with the respective contexts, even though these Old displays were repeated as before (Chun & Jiang, 1998 exp. 3). Thus, in order for learning to occur, the contexts must be predictive of target location, serving to reduce uncertainty (E.J. Gibson, 1966).

Hence, implicit learning and memory of predictive visual context guides spatial attention. Importantly, contextual cueing can be obtained from visual contexts defined by other types of invariant information. For example, visual contexts can be defined by the semantic identity of the objects comprising a display rather than their spatial layout. In one experiment using novel shapes, faster search was shown for targets whose identities covaried with the identities of contextual distractors (Chun & Jiang, 1999). This semantic contextual cueing of target identity occurred even for targets that appeared in variable locations amidst variable distractor configurations. Thus, contextual cueing is not limited to facilitation of target locations, and contextual learning is not limited to contexts defined by spatial layout. In fact, even dynamic visual displays can be learned. Consider what a football quarterback views when trying to find a receiver weaving through a sea of moving players. This situation can be mimicked in a search task in which the target and distractors all move around the screen amidst each other in random directions. Chun and Jiang (1999) showed that search time is facilitated if the target motion trajectory is correlated with a repeated set of distractor motion trajectories. Such dynamic regularities are important for a variety of situations such as driving or team sports such as basketball or soccer. Thus, these findings show that attention can be directed by invariants afforded by a wide variety of complex, global image properties.

Contextual cueing has been investigated in lab tasks using the visual search paradigm, but we believe it is highly relevant to the study of natural scene processing and change blindness. We first note that similar contextual cueing effects have been observed in other tasks using natural scenes. Testing Rhesus monkeys, Sheinberg and Logothetis (1998) demonstrated that repeated exposure to complex natural scenes produced faster search response times and fewer wasteful eye movements during search for embedded target objects. And in humans, similar benefits were observed for repeated natural scene images (reviewed in Cohen, Poldrack, & Eichenbaum, 1997). Second, applicability to change blindness paradigms seems likely also. Target locations in search tasks may correspond to regions of interest in real scenes. Rensink et al. (1997) demonstrated that changes occurring in “regions of interest” of natural scenes (such as the changing position of a helicopter viewed through a cockpit window) are noticed more readily than changes occurring in aspects of a scene less relevant to a particular context (such as shadows). Namely, contextual cueing operates to prioritize context- and task-relevant aspects of complex images over others. This is achieved through associative learning between the context and task-relevant variables that facilitate behaviour. Also, because natural scenes may typically contain several “regions of interest”, it is interesting to note that context can prioritize more than one target location in lab-based contextual cueing tasks (albeit with smaller cueing effects given a fixed amount of experience, see Chun & Jiang, 1998, exp. 6).

Thus, implicit memory mechanisms serve to facilitate a perceiver's ability to extract information from complex images. In everyday perceptual interactions in the real world, contextual cueing mechanisms may come into play in several ways. First, contextual cueing may facilitate interactions with a given scene in the environment, sampled intermittently as one looks away and back over short time scales of a few seconds. Second, contextual cueing can facilitate interactions with commonly viewed scenes such as one's office desk, one's home, or one's working environment, experienced day by day, week by week. Unpublished experiments suggest that contextual cueing effects can tolerate a significant amount of variability, as long as task-relevant, key features remain invariant. In addition, implicit memory for contextual information persists for at least a day, and we are testing the likelihood that it persists for longer intervals. Finally, we are exploring the possibility that visual learning mechanisms allow perceivers to develop more generic, abstract perceptual schemas that can be applied to different scenes (exemplars) of the same prototypical context (e.g. such as kitchen or office). Visual processes are highly sensitive to how objects covary with each other over visual experience (Chun & Jiang, 1999), so it seems likely that visual learning mechanisms will guide visual processing at such abstract levels also.

In summary, observers implicitly encode the visual context of targets whenever that contextual information is invariant over time and covaries with a variable important for visual behaviour, such as spatial location or object identity in search. In subsequent encounters with previously viewed scenes, this implicit contextual information is activated quickly and automatically to guide the deployment of attention to task-relevant aspects of complex images.

CONCLUSIONS

Work in change blindness and visual integration have revealed that the content of visual memory is remarkably impoverished, at least when its capacity is assessed using explicit report procedures (Grimes, 1996; McConkie & Currie, 1996; Rensink et al., 1997; Simons & Levin, 1997). The visual system appears to minimize the amount of information it needs to encode (Hayhoe, this issue). As such, the visual world serves as its own memory (Dennett, 1991; Horowitz & Wolfe, 1998; MacKay, 1973; O'Regan, 1992; Rensink et al., 1997; Simons & Levin, 1997). We support this view, but our goal here is to provide an important qualification.

The lack of awareness of details across views does not imply an absence of information persistence, and we propose that various mechanisms allow specific visual information to be preserved across repeated views of scenes,

biasing the deployment of visual attention. Implicit memory traces, not available to conscious awareness, are laid down during visual processing. Visual processing benefits from the accumulation of information provided by the spatial and temporal context of past views. This provides fine-tuned, internal indexing to the visual memory embodied in the external environment.

This paper discussed two independent mechanisms that provide such guidance. In priming of popout, implicit traces of attended features or locations facilitate processing of the same features or locations. In contextual cueing, the predictive context of targets is implicitly encoded and used to guide attention towards targets when this context is re-encountered in future perceptual episodes. Although priming of popout and contextual cueing are distinct, independent mechanisms, what's common is that both of these operate in an implicit manner to guide attention. And both mechanisms rely on specific memory traces of past experience, providing more precise control than "gist" alone. Thus, in everyday visual processing, contextual cueing mechanisms may serve to initiate regions of interest within a specific, complex scene which was previously viewed a few saccades, a few minutes, or perhaps even a few days in the past. The short-term memory of PoP could then ensure that fixations of functional value would be continued without supervision. Then when new regions of interest are selected, this process could begin anew, helping to guide fixations in the moment. These memory-based mechanisms of PoP and contextual cueing interact with other visual mechanisms of attentional guidance.

Altogether, these findings highlight the need to consider memory mechanisms in vision (Chun & Jiang, 1998; Desimone & Duncan, 1995; Ward, 1998). Memory biases what objects in the visual world should be attended. Implicit memory mechanisms allow attentional allocation to benefit from perceptual experience, allowing for efficient guidance and deictic indexing, and compensating for the ephemeral nature of visual information across views. The advantages of implicit representations are that they release effortful, conscious processes for other tasks (Lewicki et al., 1992), and implicit representations are generally more robust and resistant to decay or interference (Reber, 1989). Most important, the capacity of implicit representations is much higher than that accessible to explicit retrieval operations needed for change blindness tasks (see also Fernandez-Duque & Thornton, this issue).

In closing, accepting the view that minimal information from a scene is consciously retained in the head at any given time, we focused on mechanisms that may ease the transition from one view to the other. In addition to generic gist and limited-capacity transsaccadic memory, we argue that instance-based memory traces of past, perceptual interactions are retained to guide attention and action in subsequent encounters. A detailed scene serves as its own memory, but the visual system employs several strategies to provide temporal continuity and efficient access to a scene's details.

REFERENCES

- Averbach, E., & Coriell, A.S. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*, 309–328.
- Bacon, W.F., & Egeth, H.E. (1994). Overriding stimulus-driven attentional capture. *Perception and Psychophysics*, *55*, 485–496.
- Ballard, D.H., Hayhoe, M.M., & Pelz, J.B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*, 66–80.
- Bichot, N., & Schall, J.D. (1999). Saccadic target selection in macaque during feature and conjunction visual search. *Visual Neuroscience*, *16*, 81–89.
- Biederman, I., & Cooper, E.E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 121–133.
- Biederman, I., & Gerhardstein, P.C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.
- Biederman, I., Mezzanotte, R.J., & Rabinowitz, J.C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177.
- Bravo, M., & Nakayama, K. (1992). The role of attention in different visual search tasks. *Perception and Psychophysics*, *51*, 465–472.
- Broadbent, D.E. (1958). *Perception and communication*. London: Pergamon Press.
- Carlson-Radvansky, L.A., & Irwin, D.E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1441–1458.
- Chun, M.M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.
- Chun, M.M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, *10*, 360–365.
- Cohen, N.J., Poldrack, R.A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory*, *5*, 131–178.
- Dennett, D.C. (1991). *Consciousness explained*. Boston: Little, Brown & Co.
- DeSchepper, B., & Treisman, A. (1996). Visual memory for novel shapes: Implicit coding without attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 27–47.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Fernandez-Duque, D., & Thornton, I.M. (this issue). Change detection without awareness: Do explicit reports underestimate the representation of change in the visual system? *Visual Cognition*, *7*, 323–344.
- Fischer, B. (1987). The preparation of visually guided saccades. *Reviews of Physiology Biochemistry and Pharmacology*, *106*, 1–35.
- Folk, C.L., Remington, R.W., & Johnston, J.C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030–1044.
- Found, A., & Müller, H.J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a “dimension-weighting” account. *Perception and Psychophysics*, *58*, 88–101.
- Gibson, E.J. (1966). *Perceptual development and the reduction of uncertainty*. Paper presented at the 18th International Congress of Psychology, Moscow.

- Gibson, E.J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.
- Gibson, J.J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In A.A. Kathleen (Ed.), *Vancouver studies in cognitive science, Perception Vol. 5*. (pp. 89–110). New York: Oxford University Press.
- Hayhoe, M. (this issue). Vision using routines: A functional account of vision. *Visual Cognition*, 7, 43–64.
- Henderson, J.M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, 8, 51–55.
- Henderson, J.M., & Anes, M.D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 826–839.
- Henderson, J.M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception and Psychophysics*, 45, 196–208.
- Hoffman, J.E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics*, 57, 787–795.
- Horowitz, T.S., & Wolfe, J.M. (1998). Visual search has no memory. *Nature*, 394, 575–577.
- Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 604–610.
- Irwin, D.E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, 23, 420–456.
- Irwin, D.E. (1992a). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 307–317.
- Irwin, D.E. (1992b). Perceiving an integrated visual world. In D.E. Meyer & S. Kornblum (Eds), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 121–142). Cambridge, MA: MIT Press.
- Irwin, D.E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, 5, 94–100.
- Kaheman, D., Treisman, A., & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature*, 334, 430–431.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35, 1897–1916.
- Levin, D.T., & Simons, D.J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review*, 4, 501–506.
- Lewicki, P., Hill, T., Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist*, 47, 796–801.
- Logan, G.D. (1988). Towards an instance theory of automatization. *Psychological Review*, 95, 492–527.
- Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- MacKay, D.M. (1973). Visual stability and voluntary eye movements. In R. Jung (Ed.), *Handbook of sensory physiology* (Vol. 8/3, pp. 307–331). Berlin: Springer-Verlag.
- Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory and Cognition*, 22, 657–672.
- Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: II. The role of position. *Perception and Psychophysics*, 58, 977–991.

- Maljkovic, V., & Nakayama, K. (in press). Priming of pop-out: III. A short-term implicit memory system beneficial for rapid target selection. *Visual Cognition*.
- McConkie, G.W., & Currie, C.B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 563–581.
- McConkie, G.W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R.B. Ruddell (Eds), *Theoretical models and processes of reading* (pp. 137–162). Newark, DE: International Reading Association.
- McPeck, R.M., Maljkovic, V., & Nakayama, K. (in press). Saccades require focal attention and are facilitated by a short-term memory system. *Vision Research*.
- Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 411–422). Cambridge, UK: Cambridge University Press.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- 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.
- O'Regan, J.K., & Levy-Schoen, A. (1983). Integrating visual information from successive fixations: Does trans-saccadic fusion exist? *Vision Research*, *23*, 765–768.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, *44*, 369–378.
- Phillips, W.A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics*, *16*, 283–290.
- Pollatsek, A., Rayner, K., & Collins, W.E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, *113*, 426–442.
- Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.
- Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D.G. Bouwhuis (Eds), *Attention and performance X* (pp. 55–66). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Potter, M.C. (1975). Meaning in visual search. *Science*, *187*, 965–966.
- Pylshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*, 65–97.
- Rayner, K., McConkie, G.W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, *12*, 206–226.
- Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235.
- Rensink, R.A. (this issue). The dynamic representation of scenes. *Visual Cognition*, *7*, 17–42.
- Rensink, R.A., O'Regan, J.K., & Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368–373.
- Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–518.
- Sheinberg, D.L., & Logothetis, N.K. (1998). Implicit memory for scenes guides visual exploration in monkey. *Society for Neuroscience Abstracts*, Vol. 24, Part 2, 1506.
- Shepherd, M., Findlay, J.M., & Hockey, R.J. (1986). The relationship between eye movements and spatial attention. *Quarterly Journal of Experimental Psychology A*, *38*, 475–491.
- Simons, D.J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, *7*, 301–305.
- Simons, D.J., & Levin, D. (1997). Change blindness. *Trends in Cognitive Science*, *1*, 261–267.

- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74, 1–29.
- Stadler, M.A., & Frensch, P.A. (Eds.). (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage Publications.
- Tarr, M.J., & Bühlhoff, H.H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1494–1505.
- Treisman, A. (1993). The perception of features and objects. In L. Weiskrantz & A.D. Baddeley, (Eds), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent* (pp. 5–35). Oxford, UK: Clarendon Press/Oxford University Press.
- Tulving, E., & Schacter, D.L. (1990). Priming and human memory systems. *Science*, 247, 301–306.
- Ward, R. (1998). Vision in the eternal present. *Nature*, 394, 519.
- Wolfe, J.M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). Hove, UK: Psychology Press.
- Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223–256). Hove, UK: Psychology Press.
- Yarbus, A.L. (1967). *Eye movements and vision*. New York: Plenum.