Current Directions in Psychological Science

http://cdp.sagepub.com/

Representations in Visual Cognition : It's About Time

Wieske van Zoest, Amelia R. Hunt and Alan Kingstone

Current Directions in Psychological Science 2010 19: 116
DOI: 10.1177/0963721410363895

The online version of this article can be found at:
http://cdp.sagepub.com/content/19/2/116

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE
Association for Psychological Science

Additional services and information for *Current Directions in Psychological Science* can be found at:

Email Alerts: http://cdp.sagepub.com/cgi/alerts

Subscriptions: http://cdp.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Downloaded from cdp.sagepub.com at University of Aberdeen on May 18, 2010

ASSOCIATION FOR
PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
19(2) 116-120
© The Author(s) 2010
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0963721410363895
http://cdps.sagepub.com

Representations in Visual Cognition: It's About Time

Wieske van Zoest[1], Amelia R. Hunt[2], and Alan Kingstone[3]
[1] Vrije Universiteit Amsterdam, [2] University of Aberdeen, and [3] University of British Columbia

Abstract

Visual cognition relies on changing representations of visual information. The dynamic nature of representations is demonstrated by new findings in human perception and attention showing that the influence of specific aspects of a stimulus on behavior changes dramatically over time. As a consequence, performance depends on what point in time responses are measured. Specifically, quick responses to early representations of a new scene are heavily influenced by the relative salience of different stimuli in the visual field. Slow responses based on later representations of the same scene are informed by more complex information that integrates prior knowledge and observer goals. Thus, as a result of the dynamic nature of representations, the kind of information that is prioritized depends on the moment in time the representation is accessed. Examining how representations change over time can lead to new and important insights in a wide range of domains of human cognition.

Keywords

attention and perception, dynamics, visual representations

Time is critically important in our experience of the visual world. Take for example the painting *Old Woman at Prayer*, by Nicholas Maes (Fig. 1). You will likely notice the woman's face and hands before other elements in the scene. While you might describe this as a peaceful scene after a brief glance, the longer you look, the more precarious it becomes (e.g., the cat about to move on the salmon, the knife and bread sitting precariously on the edge of the table, delicate crockery right below a heavy book that is held back unsteadily by a funnel). This painting illustrates that the way information in a visual scene is perceived and that information's influence on behavior depends on how much time an observer has spent encoding it.

Dynamic Effects of Stimulus Salience

Evidence for this perspective is found in recent studies of attention and eye movements (Hunt, von Mühlenen, & Kingstone, 2007; van Zoest, Donk, & Theeuwes, 2004). These studies used a modified visual-search task called the additional-singleton paradigm (e.g., Theeuwes, 1992). In this task, observers must locate a particular target item that is presented among a number of uniform nontargets. On some trials, an additional item may be presented that is different from both the target and the nontargets; this irrelevant item is referred to as the singleton distractor. For example, in the search displays in Figure 2, participants were to locate the element tilted to the right, which was presented among vertically oriented nontargets. The additional-singleton distractor was a line element tilted to the left. Results have shown that the singleton distractor may interrupt search for the target (Bacon & Egeth, 1994; Theeuwes, 1992). That a completely irrelevant but salient object may hamper search suggests that visual selection is driven, at least in part, by inflexible responses to visually salient properties of the environment. However, the singleton distractors do not always interrupt search (Bacon & Egeth, 1994). The ability to ignore irrelevant distractors suggests that visual selection may be guided by the intentions and strategies of the observer. Together these results may appear to be at odds, but they can be reconciled by the general conclusion that visual search is controlled sometimes by low-level stimulus properties and at other times by the goals and intentions of the observer.

What determines whether visual selection is guided by stimulus properties or by the observer's goals? The timing of selection seems to be a critical factor. Evidence suggests that observers are more prone to interference from irrelevant distractors when they are quick to respond (e.g., Hunt et al.,

Corresponding Author:
Wieske van Zoest, Department of Cognitive Psychology, Vrije Universiteit Amsterdam, van der Boechorststraat 1, 1081 BT Amsterdam, the Netherlands
E-mail: w.van.zoest@psy.vu.nl

Downloaded from cdp.sagepub.com at University of Aberdeen on May 18, 2010

Fig. 1. The painting *Old Woman at Prayer*, also known as *Prayer Without End*, by Nicholas Maes. It was completed in 1656 and is currently held at the Rijksmuseum in Amsterdam, the Netherlands.

2007). Although this finding suggests people may simply be sacrificing accuracy for speed, there is more to it than that. The early effect of the distractor seems to be driven more generally by stimulus salience (Donk & van Zoest, 2008; van Zoest et al., 2004). Independent of whether an object is defined as a distractor or target, the most visually salient objects are selected when observers respond quickly. For example, in one experiment (van Zoest et al., 2004) observers were instructed to make an eye movement (called a saccade) to the target item; the salience of an additional singleton distractor was varied such that the irrelevant singleton could be less, equally, or more salient than the target (see Fig. 2). Results showed that when observers started their eye movements soon after the display appeared, gaze tended to be directed toward the most salient object presented in the display: Observers were very accurate when the most salient item in the display happened to be the target, but they made many incorrect saccades toward a distractor when it was the most salient item. As saccadic latency (time between presentation of the stimulus and movement of the eyes) increased, the influence of stimulus salience decreased, with no effect of stimulus salience at the slowest saccadic latencies. These slower saccades were driven solely by the identity of the target. Note that this critical interaction of visual selection with response time was revealed only when quick and slow responses were analyzed separately.

Further investigations revealed that stimulus representations change regardless of whether voluntary processes are on line to guide identity-driven selection (Donk & van Zoest, 2008). When participants were instructed to locate the item that was most visually salient, they were able to do so only when they were quick to respond—that is, even though stimulus salience was explicitly task relevant, observers were unable to guide slower responses to the most salient item. This result suggests that the relative representation of salience degrades over time. If information about target salience had persisted in the visual system, correct selection of the salient location should have been possible regardless of whether responses were quick or slow.

Dynamic Representations Across Modalities

The finding that responses made earlier in time are influenced by different factors than responses to the same displays made later in time poses potential challenges for comparing responses from different modalities. Manual responses tend to be hundreds of milliseconds slower than eye-movement responses. Indeed, given the large differences in behavioral performance between modalities, some have argued that these different responses rely on their own response-specific representations (Aglioti, DeSouza, & Goodale, 1995). To investigate to what extent different modalities rely on a shared representation, Hunt et al. (2007) examined the time courses of saccadic eye movement and manual joystick responses to targets in the presence of sudden-onset distractors. Observers were instructed to locate a color singleton target either by making a saccadic eye movement or by manually moving a joystick in the direction of this item. Eye movements, but not joystick movements, were captured by the onset distractor and incorrectly moved to this location before moving on to the target location. However, when participants were instructed to make very fast joystick responses, thus bringing the reaction-time distributions for the two response modalities closer together, the differences in the effect of the distractor across the two modalities were profoundly reduced: Fast eye and manual responses were directed toward the onset distractor, and slower eye and manual responses were directed toward the target. As a similar pattern of results across time was found in both response modalities, it was concluded that the distractor is influencing a representation of the visual array that is shared across response systems (Hunt et al., 2007). Responses from different modalities access the representation of the visual array at different moments in time.

Dynamics in the Brain

Neurophysiological correlates lend further support to the idea that time is critically important to representations of the visual world. A recent study that looked at event-related potentials (ERP) in the brain (Hickey, van Zoest, & Theeuwes, in press) found that the time course observed in studies of overt selection (Hunt et al., 2007; van Zoest et al., 2004) may reflect a pattern of covert selection (stimulus selection without any concomitant movement of the eye or head). As in previously discussed

Downloaded from cdp.sagepub.com at University of Aberdeen on May 18, 2010

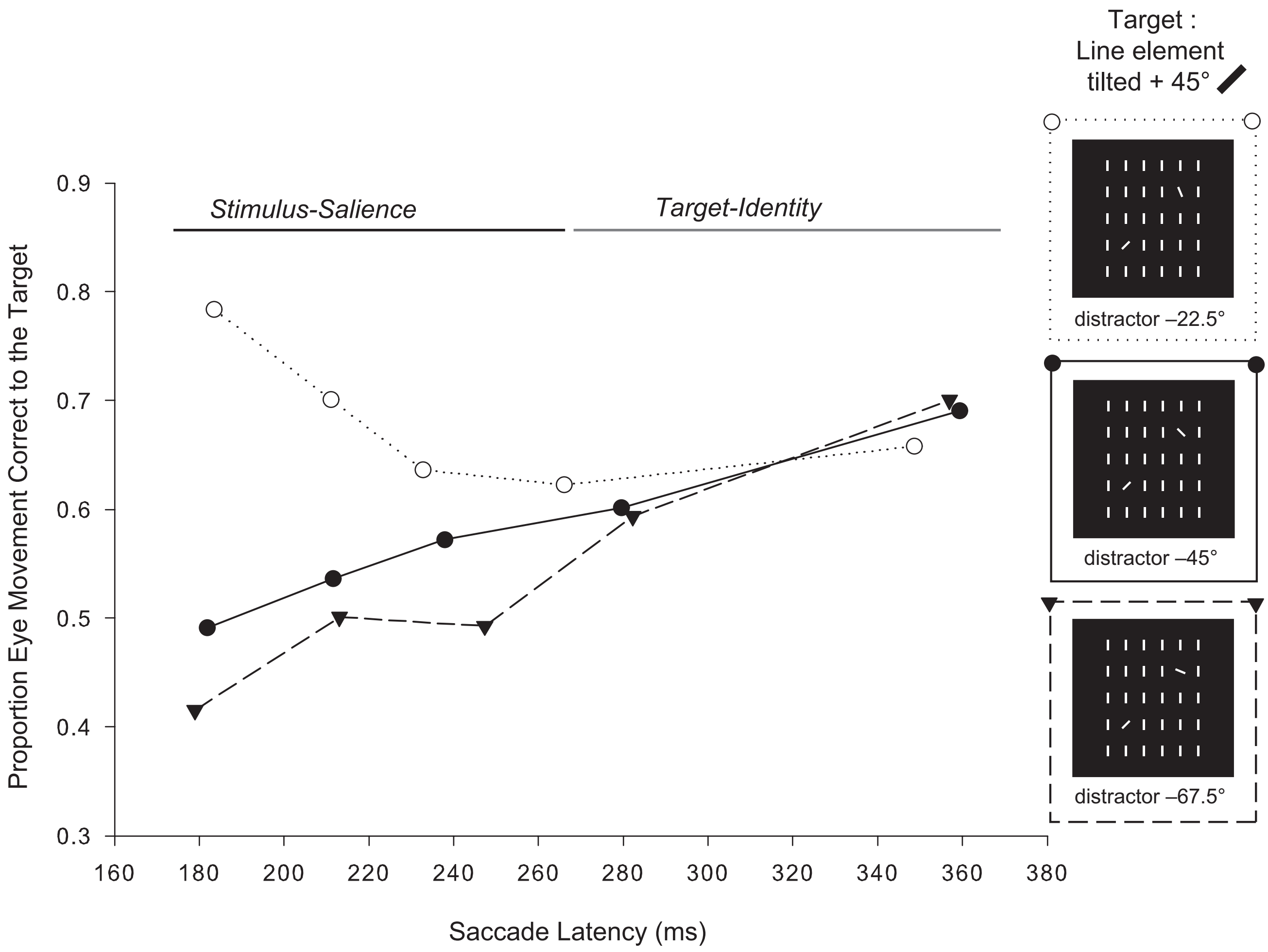

Fig. 2. Proportion of correct eye movements to a target as a function of saccade latency, with distractors of different degrees of salience (van Zoest et al., 2004, Experiment 4). Participants were required to make an eye movement to a target line element tilted 45° that was presented among a uniform series of vertical elements as well as one additional item, the distractor. The irrelevant distractor could be less salient than the target (-22.5°: white circles, dotted line), equally salient (-45°: black circles, solid line) or more salient than the target (-67.5°: black triangles, dashed line). The results demonstrated that quick eye movements were directed to salient items, independently of the identity of the target; no effect of stimulus salience was found for eye movements executed later in time.

studies, Hickey et al. presented participants with the additional-singleton paradigm and investigated the time course of responses. The results demonstrated that rapid deployment of attention, as reflected in early onset of the ERP component called the N2pc, was associated with increased distractor processing and with an increased response time to identify the target. Slower deployment of attention, as reflected by later onset of the N2pc, was associated with increased target processing and a decreased response time to identify the target. In other words, when the representation was probed early in time following stimulus presentation, the distractor was selected; when the representation was probed later in time, the target was selected for further processing.

Direct recordings from the primary visual cortex (area V1) in monkeys demonstrate the critical role of time in the representation of visual information at a neuronal level (Roelfsema, Tolboom, & Khayat, 2007). Monkeys were trained to segregate two figures from the background and to select one of the figures by means of an eye movement. This task comprised a number of different processes, including the processing of features and figures as well as the selective processing of the target object by means of attention. These different processes are often thought to be represented at different levels in the visual cortical hierarchy, with features being represented at low-level visual areas, figures at intermediate levels, and attentional processes at higher cortical areas (e.g., Corbetta, Patel, & Shulman, 2008). However, Roelfsema et al. (2007) showed that the same V1 neurons carry signals related to all of these processes; the critical point being that these different processes correspond to different moments in time after the presentation of the visual image. The initial neuronal responses triggered by the onset of the visual stimulus convey information about the features, while figures are represented by these same neurons only after a small delay. Attentional processes are represented by the same neurons after yet another delay. Thus,

Downloaded from cdp.sagepub.com at University of Aberdeen on May 18, 2010

instead of having separate representations at different levels of the hierarchy, representational dynamics allow for the same cells to serve different functions at different stages of processing; this is also known as multiplexing (Di Lollo, Enns, & Rensink, 2000).

The idea of dynamic visual representations is in line with theories of information processing that have emphasized the importance of dynamics in feedforward and recurrent processing in the brain (Di Lollo et al., 2000; Lamme & Roelfsema, 2000). For example, according to Lamme and Roelfsema, processing that occurs immediately following the presentation of an image relies on rapid feedforward connections that spread from low-level to high-level areas. During later processing, information from horizontal or feedback connections is incorporated. Whereas initial responses reflect sensory processing by the feedforward sweep, responses at longer latencies correlate with higher-level information, such as prior knowledge and observer goals (and in the context of Lamme and Roelfsema's argument, with conscious perception of the stimulus). In this sense, initial representations that strongly signal stimulus salience may become more tailored to the viewer's goals and intentions as salience degrades and re-entrant signals converge on the representation.

Time as an Independent Variable

The idea that time is critically important in visual cognition is not new. Many studies use time as an independent variable to investigate how processing changes. In these methods, two events (i.e., prime and target) are presented in sequence and observers are required to respond to the second event. How the response to the second event is influenced by the first event depends on the amount of time between them. Results typically reveal that performance benefits occur when a brief time separates the two events; in contrast, performance costs occur when the time between the events is longer. For example, when a sudden event draws spatial attention to a location automatically, processing at that location is initially facilitated. Later, processing at the same location is inhibited. This is referred to as inhibition of return and is interpreted as a bias against returning to a location that has just been attended (Posner & Cohen, 1984). Early-facilitation and later-inhibition effects also characterize response priming to unseen (subliminal) primes. When the interval between a prime and the target is short, positive priming occurs, such that responses are faster if the prime is visually similar to the target. When the interval between the prime and target is longer, negative priming occurs, such that responses are paradoxically slower when the prime is visually similar to the target (Lleras & Enns, 2004). Similar biphasic priming effects are found in lexical-decision tasks, in which participants are explicitly instructed to expect a target word that is of a different semantic category than the prime. When the interval between a prime word and target word is short, rapid automatic facilitation of the target word is found when this word is semantically related to the prime. When the time between the two events is longer, responses to target words that are semantically related are inhibited in order to facilitate processing of target words that are of the expected category. These results demonstrate that higher-level expectations regarding the upcoming target affect performance only later in time (Neely, 1977).

Though these studies clearly reveal an important role for time in processing dynamics, a potential pitfall of this approach is that the role of time is measured in discrete conditions. An indirect and unintentional consequence of this approach is that when differences are found across discrete time conditions, they tend to be ascribed to discrete processing mechanisms such as facilitation and inhibition. This interpretation carries the assumption that the represented information is constant but that processes using and accessing that representation vary over time. However, these differences in performance and processing may instead be explained by a single process that operates on information that changes over time.

Conclusions and Further Directions

Information develops over time, and the quality of information changes. The stimulus representation acted on at one point in time is not the same as the stimulus representation that is acted on at a different point in time. Dynamic representations pose a serious challenge for comparing different tasks that demand different amounts of information for successful performance. For example, visual search for a unique feature is quick in comparison to search for a conjunction of features. Though feature search is often driven by saliency, conjunction search is driven by specific knowledge regarding the identity of a target. However, these differences may in part be caused by the way representations change over time. Pop-out visual search is never slow and conjunction search is rarely quick. If it were possible to more closely align the different time courses of feature and conjunction search, one may find that they are driven by a common representation that changes with time.

In a similar vein, irrelevant distractors interrupt visual processing primarily when displays are homogeneous and "perceptual load" is low. More heterogeneous displays in which the perceptual load is high are typically resistant to the effect of irrelevant distractors. Comparing these two tasks directly is potentially problematic, as participants are faster to respond to conditions of low load than they are to conditions of high load. In turn, responses in low-load conditions may be accessing visual representations at an earlier point in time than they do in conditions of high load. Given that the effects of stimulus salience are time dependent and primarily affect early responses to a representation, difficult tasks that take more time may not be affected by irrelevant distractors simply because the representation of stimulus salience has changed. Because representations of information are not static, it is essential to consider when a response is being executed. Many aspects of visual cognition may function through changes of representation over time.

Appreciation of this point opens the door for future investigations in at least two important ways. At a local empirical level, it provides researchers with the opportunity to gain traction on problems that have often bedeviled scientific analysis. For instance,

Downloaded from cdp.sagepub.com at University of Aberdeen on May 18, 2010

much of the variance between participants might be reconciled by an appreciation for the fact that different observers are responding at different times, and therefore different observers may quite literally be acting on different qualitative forms of information. Similarly, differences between response modalities—or sensory modalities for that matter (Kingstone, 1992)—which heretofore have been thought to reflect modality-specific processing, may actually reflect the fact that different modalities engender different effects because they produce substantially different response latencies. If the times across responses were equated, the differences between the responses may also converge. Importantly, if they do not converge, then researchers will have evidence that the representations that are being acted on are qualitatively different.

At a broader theoretical level, it is clear that the models of cognition that researchers develop will need to be dynamic in nature even when they are constrained to the investigation of human performance by a series of stimulus trials. Within each trial, the role of time cannot be overlooked. In building dynamic models that can explain human processing on a timescale of 1 to 2 seconds, the hope is that we can at last answer the challenge of scientists such as J.J. Gibson, who emphasized the idea that humans exist within a continuous dynamic sea of everchanging information (e.g., Gibson, 1958/2009), a sea that cannot be understood by studying performance in discrete individual stimulus trials.

Recommended Reading

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*, 791–804. Review discussing how feedforward and feedback processing relates to implicit and explicit forms of visual perception, and discusses how this may explain various phenomena in visual cognition.

Roelfsema, P.R., Tolboom, M., & Khayat, P.S. (2007). (See References). Provides neurophysiological evidence for the idea that different psychological processing stages map onto distinct time episodes in the visual cortex.

VanRullen, R. (2003). Visual saliency and spike timing in the ventral visual pathway. *Journal of Physiology-Paris*, *97*, 365–377. Paper discussing various views on visual salience and proposes that spike latency plays an important role in the encoding of visual salience.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

W.v.Z. was supported by a grant from the Netherlands Organization for Scientific Research.

References

Aglioti, S., DeSouza, J.F., & Goodale, M.A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology*, *5*, 679–685.

Bacon, W.F., & Egeth, H.E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*, 485–496.

Corbetta, M., Patel, G., & Shulman, G.L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, *58*, 306–324.

Di Lollo, V., Enns, J.T., & Rensink, R.A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, *129*, 481–507.

Donk, M., & van Zoest, W. (2008). Effects of salience are short-lived. *Psychological Science*, *19*, 733–739.

Gibson, J.J. (2009). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology, 100*, 259–271. (Original work published 1958).

Hickey, C., van Zoest, W., & Theeuwes, J. (in press). The time course of exogenous and endogenous control of covert attention. *Experimental Brain Research*.

Hunt, A.R., von Mühlenen, A., & Kingstone, A. (2007). The time course of attentional and oculomotor capture reveals a common cause. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 271–284.

Kingstone, A. (1992). Combining expectancies. *The Quarterly Journal of Experimental Psychology*, *44A*, 69–104.

Lamme, V.A., & Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience*, *23*, 571–579.

Lleras, A., & Enns, J.T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *Journal of Experimental Psychology: General*, *133*, 475–493.

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory—roles of inhibition-less spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouman & D.G. Bouwhuis (Eds.), *Attention and Performance X.* (pp. 531–556). Hillsdale, NJ: Erlbaum.

Roelfsema, P.R., Tolboom, M., & Khayat, P.S. (2007). Different processing phases for features, figures, and selective attention in the primary visual cortex. *Neuron*, *56*, 785–792.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606.

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 746–759.

Downloaded from cdp.sagepub.com at University of Aberdeen on May 18, 2010