terials is associated specifically with changes in cortical visual processing regions (Schacter & Buckner 1998). Prior exposure also shapes the nature of subsequent visual exploration of those materials when presented again, as measured by eye movements. Our own work has shown that the pattern of viewing of faces (Althoff & Cohen 1999) and of scenes (Ryan et al. 2000) is altered upon repeated presentations. These effects can occur in the absence of conscious or explicit memory for the (previously presented) primed items.

Our work has shown that visual representations include information about relations among the constituent elements of scenes, and are used by viewers to guide further visual exploration (Ryan et al. 2000). Subjects saw a series of scenes and then subsequently saw either repeated versions or manipulated versions of the scenes involving some difference in the relations among scene elements. Increased viewing was directed to the changed region in manipulated scenes, most notably when subjects were *unaware* of any scene manipulation. Accordingly, relational information was maintained in visual representations and was used by observers to guide subsequent eye movement behavior, even in the absence of conscious awareness of the change.

Chun and Jiang's (1998) findings of contextual cuing, in which subjects were faster in finding a target in a repeated display of distracters compared to a new display, as long as the position of the target within the display remained constant, lead to similar conclusions. Viewers were apparently able to form representations containing information about the target object and its relations to the distracters in the display. Viewers were then able to use this relational information to guide further search behavior even in the absence of conscious awareness of the information.

Another empirical evidence cited by **O&N** as proof against internal representations of the visual world is the phenomenon of change blindness, in which viewers are surprisingly poor at reporting changes made to scenes on-line. However, our most recent work demonstrates eye movement evidence of change *detection* (Ryan & Cohen, submitted), emphasizing the presence of visual memory representations that are maintained on-line and used to guide subsequent viewing behavior. After viewing an initial scene, subjects saw it again immediately, either repeated in its original form or in a manipulated version. Subjects directed increased viewing to the very region of change in manipulated scenes, even within the first gaze to the region. Additionally, the overall pattern of viewing of manipulated scenes was altered compared to viewing of repeated scenes, even when subjects were unaware that a change had occurred. Similar findings have also been reported by Hollingworth et al. (in press).

Taken together, these findings indicate that internal memory representations of the visual world play a critical role in vision and visual exploration. Visual representations containing information about viewed objects and the relations among them are maintained in viewers' brains and used to compare with the currently presented visual display in order to guide exploratory eye movement behavior.

Reexamining visual cognition in human infants: On the necessity of representation

Matthew Schlesinger

Department of Psychology, Southern Illinois University, Carbondale, IL 62901. **matthews@siu.edu** **http://www.siu.edu/~matthew**

Abstract: The sensorimotor account of vision proposed by O'Regan & Noë (O&N) challenges the classical view of visual cognition as a process of mentally representing the world. Many infant cognition researchers would probably disagree. I describe the surprising ability of young infants to represent and reason about the physical world, and ask how this capacity can be explained in non-representational terms. As a first step toward answering this question, I suggest that recent models of embodied cognition may help illustrate a way of escaping the representational trap.

O'Regan & Noë (**O&N**) present a compelling account of visual cognition that trades in the popular myth of vision as recovery and reconstruction of the 3D world, for a leaner, meaner alternative. I enthusiastically endorse their thesis that much of "seeing the world" can be better understood as a process of recognizing how actions in the world (e.g., eye and head movements) systematically transform sensory data over time. The theory of sensorimotor contingency reaches well beyond the domain of perception and visual consciousness, raising many important questions for other areas of research. In particular, as a developmental psychologist I would like to focus my commentary on the implications of O&N's proposal for infant cognition research.

Infants mentally represent the world. Across a wide range of knowledge domains (e.g., number, objects, causality, classes), there is now unequivocal, or at least overwhelming support for the conclusion that young infants mentally represent the world (Baillargeon 1999; Meltzoff & Moore 1998; Spelke 1998; Wynn 1992; however, see also Bogartz et al. 2000; Haith 1998; Rivera et al. 1999). For example, young infants react with surprise (i.e., increased attention) when they see two dolls placed behind a screen and then only one doll revealed when the screen is removed (Wynn 1992). Infants are similarly surprised when they see a drawbridge appear to rotate through a space that was occupied a moment before by a solid box (but not if the box is shown to be made of sponge).

It is typically assumed in these studies that infants mentally reconstruct salient features of the physical environment, forming internal copies or models of the external world that not only persist over time, but also guide infants' perceptual information-gathering. Complimenting the representational account of infants' visual cognition are a number of recent computational models that illustrate how such representations might arise during early development (Mareschal et al. 1999; Munakata et al. 1997; Simon 1998).

Infants operate on their mental representations of the world. It may not only be the case that infants mentally represent the world, but that they also seem to use these representations as detailed physical models when data from the real world are temporarily unavailable (e.g., when an object is occluded). For example, infants react with surprise at the sight of a car that reappears after its path is obstructed by a large, occluded box (Baillargeon 1986). Such findings are used to conclude that infants use mental models to actively *reason* about the physical world, by forecasting or predicting the outcomes of occluded events.

The representational account of infant visual cognition creates a difficult challenge for **O&N**'s theory. Much of what they marshal as evidence for their proposal are data and thought experiments from the here-and-now, like exploring a bottle that is in one's hands. How might we explain an infant's putative ability to hold in mind both the existence and location of an occluded object, or more impressively, to recognize when that object should obstruct the movement of another occluded object, within the formal framework of sensorimotor contingencies? It is unclear how the idea of the "world as an external representation" could account for such impressive mental skills.

A first step toward escaping the representational trap. Can the sensorimotor account of visual cognition explain the apparent representational prowess of young infants? This is a difficult question, but one that I hope **O&N** can somehow answer. However, paralleling their approach are a number of researchers (e.g., Ballard et al. 1997; Nolfi & Parisi 1999; Scheier & Pfeifer 1995; Schlesinger & Parisi 2001) who (1) advocate the notion of embodied knowledge, and (2) have investigated a variety of practical tasks that would seem to require mental or symbolic representations, but that can in fact be solved by assuming that knowledge is body-based, situated, and encoded implicitly through temporally-extended actions.

As an example of this approach, I have developed an oculomotor control model that simulates the eye movements of an infant as it tracks moving objects (Schlesinger & Barto 1999; Schlesinger & Parisi 2001). Using a reinforcement-learning algorithm, the eye-movement model quickly learns to anticipate the reappearance of an occluded object. Perhaps more impressively, like the young infants in Baillargeon's (1986) study, the model also tracks the target appropriately when its path is blocked by a hidden obstacle. Note that these prospective behaviors are achieved with only a simple associative learning mechanism, and without assuming or building in the capacity for memory or prediction.

Obviously, this eye-movement model is just one small step toward trying to explain infants' visual cognition in non-representational terms. Nevertheless, the model suggests that we can know about and act in the world, indeed, even anticipate future states of the world, without having to explicitly reference a mental model or symbolic representation. In this sense, of course, an important point of agreement between my model and **O&N**'s more general theoretical account is that *visual cognition is activity.* It is unclear to me, however, how far such a theory can go toward explaining not only the (apparent) capacity for infants to represent the world in symbolic form, but also their capacity for using those putative representations to act, judge, or reason about the world in a prospective manner.

Summary. I have highlighted here the representational account of infant cognition. My challenge to **O&N** is to continue to expand and elaborate their theory in at least two directions. First, I hope that they can flesh out the sensorimotor account of vision in enough detail to provide a compelling alternative to the idea that knowledge acquisition is simply a process of copying external reality. Second, and more generally, I encourage both authors to raise their sights, and to begin thinking beyond the question of learning and, instead, toward the lifelong process of perceptual and cognitive development.

ACKNOWLEDGMENTS
This commentary was supported by a grant from the National Institute of Child Health and Human Development (1R03 HD40780-02). I thank Michael Young for his helpful comments during its preparation.

Change blindness, Gibson, and the sensorimotor theory of vision

Brian J. Scholl[a] and Daniel J. Simons[b]
[a]*Department of Psychology, Yale University, New Haven, CT 06520-8205;*
[b]*Department of Psychology, Harvard University, Cambridge, MA 02138.*
Brian.Scholl@yale.edu dsimons@wjh.harvard.edu
http://pantheon.yale.edu/~bs265
http://www.wjh.harvard.edu/~viscog/lab/

Abstract: We suggest that the sensorimotor "theory" of vision is really an unstructured collection of separate ideas, and that much of the evidence cited in its favor at best supports only a subset of these ideas. As an example, we note that work on change blindness does not "vindicate" (or even speak to) much of the sensorimotor framework. Moreover, the ideas themselves are not always internally consistent. Finally, the proposed framework draws on ideas initially espoused by James Gibson, but does little to differentiate itself from those earlier views. For even part of this framework to become testable, it must specify which sources of evidence can support or contradict each of the component hypotheses.

On its surface, the "sensorimotor account of vision" by **O'Regan & Noë** (**O&N**) has an impressive scope: among many other things, it allegedly explains the nature of visual consciousness; how sensory modalities differ; how sensation differs from perception; how we perceive a stable world despite eye movements; why visual binding is unnecessary; and why hundreds of years of philosophical analysis of the problem of qualia can be dismissed as misguided.

How can so many phenomena fall under the explanatory scope of this single theory? One reason, we suggest, is that it is not so much a coherent theory as an unstructured collection of three interesting ideas: (1) vision is active and exploratory rather than passive; (2) knowledge of sensorimotor contingencies plays a central role in conscious vision; and (3) the visual system uses the world as an "outside memory." Although most researchers would accept the first idea, the latter two are more controversial. More importantly, these ideas are essentially unrelated: each can be selectively denied while maintaining the others.

The scope of these hypotheses and the many types of evidence alleged to support them creates a substantial problem: given that the theory is not a single, structured claim, it is unclear which of the ideas are supported by which types of evidence. The sensorimotor theory is treated by **O&N** as a coherent whole, and evidence consistent with some of the ideas is inappropriately taken to substantiate the theory in its entirety.

The role (or lack thereof) of change blindness in the sensorimotor theory. Here we focus on just one instance of this error, involving *change blindness* – the phenomenon wherein surprisingly large changes go unnoticed, even when observers are actively trying to find them (see Simons 2000a). The authors view change blindness as central to their theory: "the sensorimotor approach to vision . . . has provided the impetus for a series of surprising experiments on what has come to be known as change blindness. The robustness of these results in turn serves to vindicate the framework itself" (sect. 9). We suggest that both of these claims are mistaken, and that change blindness does not directly support the sensorimotor theory.

O&N suggest that change blindness was discovered as a direct consequence of the sensorimotor theory, or more precisely, the "world as an outside memory" claim. Although this idea did provide some of the theoretical motivation for recent work on such phenomena, the initial work on change blindness was not motivated by this issue at all. Most early work on change blindness derived from the study of visual integration and focused on the detection of changes during reading. For example, McConkie and Zola (1979) showed that observers often failed to notice when every letter on a screen changed case during a saccade. Other work on the failure to notice changes, both theoretical and empirical, similarly predated the current theory (e.g., Dennett 1991; Hochberg 1986; Phillips 1974; Stroud 1955).

The notion of using the world as an outside memory (e.g., Brooks 1991; O'Regan 1992; Stroud 1955) might explain why several forms of change blindness occur: we intuitively expect to detect such changes, perhaps on account of implicit beliefs about the capacity and fidelity of internal representations, or perhaps because of implicit expectations about the range of unusual or distinctive events that will draw our attention (e.g., Levin et al. 2000; Scholl et al., submitted). In any case, accurate change detection, when it does occur (in situations which do not induce change blindness) may be driven largely by motion transients which draw attention back to the world itself (e.g., Simons et al. 2000). Though the externalized memory hypothesis might predict change blindness, however, it is not clear that the sensorimotor hypothesis would. Sensorimotor contingencies require an internal memory from one instant to the next, because detecting contingencies depends on the ability to note how an environment changes in response to actions such as a "flick of the eyes." However, if the observer relied solely on the external world to provide their memory, then nothing would ever be seen to change across such flicks of the eyes (due to saccade-contingent change blindness). How, then, would observers learn what was stable and what was variable over time and across eye and head movements?

Thus, change blindness – including the flicker and mudsplash paradigms developed by Rensink, O'Regan, and colleagues – provides no direct support for **O&N**'s sensorimotor contingency idea. In fact, it does not even directly support the externalized memory idea. Change blindness is consistent with the idea that we lack internalized, detailed representations: in the absence of such inter-