ACKNOWLEDGMENT
I am grateful to C. B. Martin and Rodolfo Llinás for helpful discussions on perception and imagery, and to the McDonnell-Pew Program in Cognitive Neuroscience and the Social Sciences and Humanities Research Council of Canada for generous support.

## Spatiotemporal unit formation

Thomas F. Shipley
*Department of Psychology, Temple University, Philadelphia, PA 19122.*
**tshipley@astro.ocis.temple.edu**

**Abstract:** Findings in dynamic unit formation suggest that completion processes reflect the optics of our world. Dynamic unit formation may depend on patterns of motion signals that are consistent with the causes of optical changes. In addition, dynamic completion conforms to a local curvature minimization constraint. Such relational aspects of vision are important to consider in linking perceptual experience and neural activity.

Much of the discussion of filling in has focused on static displays. Perceptual unit formation is also evident in moving displays, and research on surface, boundary, and path completion of moving objects highlights the importance of ecological considerations in understanding the link between neural activity and perceptual experience. At the personal level both modal and amodal completion of objects is observed. A substantial portion of Michotte's original discussion of amodal perception included dynamic displays (Michotte et al. 1964). For example, in "tunneling," an object will be seen as moving along a continuous path behind an occluder when the visible parts of the moving object's path align spatially and temporally. Apparent motion, where an object appears to move back and forth between two locations, is an example of modal completion across time.

It might be tempting to hypothesize that neurons that respond to a moving point (and apparent motion) are the locus for the neural completion of an object's changing location; however, a consideration of the optics of our world suggests that the link between neural and personal is not so simple. Sequential luminance changes on the retina may arise from two different environmental events, the motion of an object or the sequential occlusion of multiple objects or textural elements. In the first case, the local motion signals are related directly to the motion in the world; the motion signals that occur as things appear and disappear, however, are not. Hence the latter motion signals should not serve as representations of object locations in the world. These signals are not random; their spatiotemporal pattern is lawfully related to an aspect of the occluding object – the orientation of the occluding edge (Shipley & Kellman 1994; 1997). Each pair of appearances or disappearances produces a motion signal whose direction and magnitude is a function of the orientation of the occluding edge and the relative position of the changing elements. As a consequence, any pair of motion signals could be used to recover the orientation of the occluding edge. Specifically, if two motion signals, represented by vectors, are positioned so that they have a common origin, the tips of the vectors define the local orientation of the occluding edge. Discriminating between the patterns of motion signals that accompany motion and dynamic occlusion allows the visual system to use the motion signals to identify boundaries of surfaces that are not specified by luminance, texture, or other static differences. Such boundary formation processes could aid in the identification of objects seen while moving through a cluttered world.

The distinction between local element motion and occlusion has been reported by a number of researchers who have found that one of two percepts may be experienced in displays where elements appear and disappear. Depending on stimulus conditions, such as the temporal interval between changes, observers may report either (1) an edge hiding and revealing elements or (2) motion of individual elements (Shipley & Kellman 1994; Sigman & Rock 1974; Wallach 1935/1996). A parallel dichotomy is observed in static displays where a tangent discontinuity is perceived as either a corner in the world or a consequence of partial occlusion.

There are a number of other parallels between static illusory contours and spatiotemporally defined edges. In both cases the bounded region appears to have a surface quality that differs from the surround (Cicerone et al. 1995; Cunningham et al. 1998), but a surface quality difference is not necessary to see a bounding edge (Kellman & Loukides 1987; Shipley & Kellman 1994). In addition, the filled in edge is almost always seen as smooth. Corners in illusory contours are rarely reported and when shown a spatiotemporally defined hexagon subjects frequently mistake it for a circle (Shipley & Kellman 1994). Minimization of change, both across surfaces and along edges, would appear to be a general property of filling in.

The observation of covariation in perceptual aspects of an event, as well as constraints on how filling in occurs, are relevant to issues of representations in vision. Conceiving of perceptual representations as encoding all aspects of a scene is indeed problematic. A more limited meaning of representation may be appropriate, and may help focus on the importance of relationships. Gallistel (1990) offers a definition of neural representations as events or states of the nervous system that are isomorphic to events or states of the world. This definition does not require all aspects of the world be encoded, only the ones that are critical for action. Unlike many uses of the term representation, this is purely mathematical, so it focuses on what relationships (or functions) present in the represented system are preserved in the representing system (e.g., when numbers are used to represent quantities of objects, some relations such as addition and division are preserved, whereas others, such as whether the objects are soft or hard, are not).

In building a bridge between the personal and neural, it may be useful to try to understand which mathematical relationships apply in the representing system. Considering both how the nervous system may instantiate computations (Gallistel [1990] reviews a number of cases where the nervous system appears to use vector representations and vector addition occurs as a consequence of the spatial structure of the nervous system) and the ecological relationships, which suggest that certain perceptual experiences will be coupled (Epstein [1982] provides a review of cases where two perceptual experiences are clearly linked) should be helpful.

ACKNOWLEDGMENT
This work supported in part by NSF grant SBR 9396309.

## Active vision and the basketball problem

Manish Singh and Donald D. Hoffman
*Department of Cognitive Sciences, University of California, Irvine, CA 92697-5100.* **msingh@uci.edu ddhoff@uci.edu**
**aris.ss.uci.edu/cogsci/personnel/gstudents/singh/**
**aris.ss.uci.edu/cogsci/personnel/hoffman/hoffman.html**

**Abstract:** It is fruitful to think of the representational and the organism-centered approaches as complementary levels of analysis, rather than mutually exclusive alternatives. Claims to the contrary by proponents of the organism-centered approach face what we call the "basketball problem."

The target article of Pessoa et al. nicely summarizes a long list of perceptual-completion phenomena. To this list we can add a relatively new phenomenon called "dynamic color spreading" (DCS; Cicerone & Hoffman 1991; Cicerone et al. 1995) in which apparent motion triggers human vision to construct illusory surfaces and contours. Figure 1a shows one frame from a DCS movie, consisting of a square containing randomly placed dots. Most dots are red, but a few (which fall within a virtual disk) are green. Figure

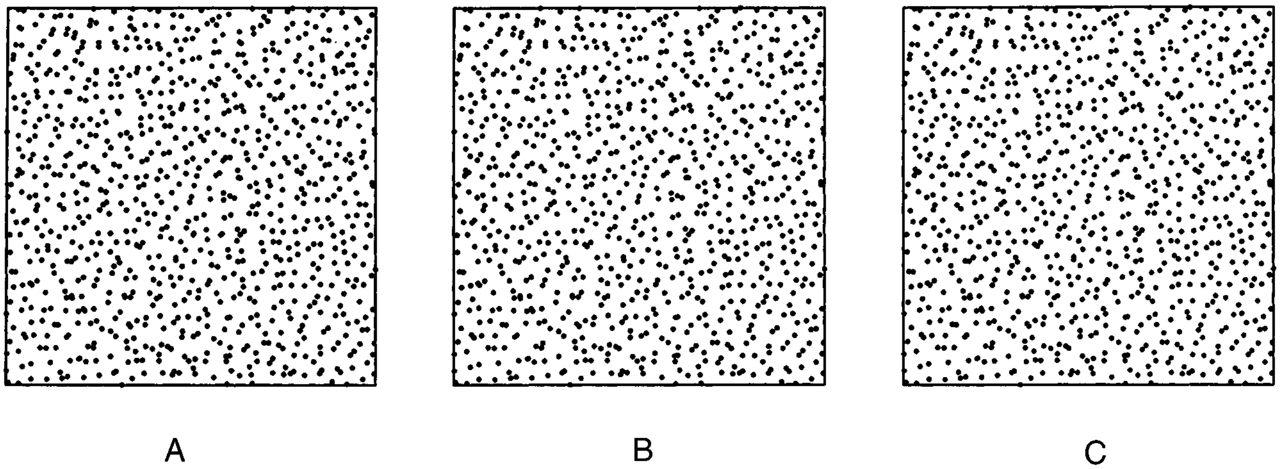


Figure 1 (Singh & Hoffman). Three frames from a dynamic color-spreading movie.

1b and 1c show two more frames from the same movie. From frame to frame no dots move – but different dots are colored green. Setting these frames in motion triggers the modal perception of a circular and transparent green filter gliding over the red dots. It can also trigger the amodal perception of an opaque green disk gliding behind the red dots. DCS is unique in that the completion of the surface and bounding contour is triggered by apparent motion.

The target article properly warns readers of the dangers of analytic isomorphism and Cartesian materialism, and argues that neural filling in need not imply either. However, its positive contribution to methodology in visual science is less clear. The authors contrast the representational approach to vision with an animal-based or activity-based approach – one in which the focus of study is not internal processing, but the whole animal interacting with its environment – and they suggest that visual science should reject the representational approach in favor of the animal-based approach.

It is not clear why these two approaches must oppose each other. No one denies the importance of understanding vision and cognition within the larger context of an organism interacting with its environment. Indeed, representational questions are often motivated by observations at the organism/personal level: given that an organism can do such and such, what computations and representations might subserve this ability? For example, at the organism level, there is indeed a "perception-action system that enables the animal to visually guide its activity and thereby visually explore its environment." But one needs to ask: What mechanisms make such a perception-action system possible? This is the kind of question that the representational approach addresses. Without answers to such questions, our understanding is bound to remain incomplete. Thus the representational and the organism-centered approaches are more fruitfully thought of as complementary levels of analysis, rather than mutually exclusive alternatives.

Pessoa et al. worry that current visual research focuses on representations and ignores issues at the level of the organism. But this research simply reflects the normal modus operandi of science: It is easier to study small pieces of the puzzle first, and complex interactions later. In this regard, it is surprising that the target article offers no systematic plan for research at the organism level. In what ways would such research differ from current psychophysics? What experiments would we run? What insights might we expect? And how would these insights obviate the need for representational accounts? If the authors wish to advocate an alternative approach for vision research, the burden is on them to detail a concrete plan of research, and show why it might be superior.

This deficit is all the more acute because Pessoa et al. do not make clear whether, by perception at the level of the organism, they mean (a) perception from a third-person perspective in which we as scientists look at the behavior of an organism and decide that it perceives something, or (b) perception from a first-person "phenomenological" perspective. For example, do they believe that the same methodology would apply in both cases?

Pessoa et al. also claim that the subpersonal account of vision creates conceptual confusions regarding what vision is, and that it is guilty of the homunculus fallacy. Again, no confusion results if one keeps clearly in mind that one is dealing with two different levels of analysis. One may talk either of an organism perceiving and acting in an environment in appropriate ways, or of internal processing that allows for such behavior. With this distinction in mind, it involves no confusion to say that the construction of certain representations is a necessary condition for the organism to see (Hoffman 1998; Singh & Hoffman 1997). Take away those representations and the organism no longer sees. Nor is a homunculus fallacy involved: the representations that are constructed need not be passed on to other "higher centers" for further interpretation – some representations are by themselves sufficient to trigger appropriate responses by the organism. Hence the representational approach to vision in itself entails neither Cartesian materialism, nor analytic isomorphism, nor the homunculus fallacy.

Even if we agree, for the sake of argument, that the need for representations is minimized through reliance on perceptually guided tasks of the animal as a whole, it is nevertheless true that the animal engages in many such tasks, and that it must therefore use many representations. We do not agree, however, that this need is minimized through reliance on perceptually guided tasks. Consider the perceptually guided task of playing professional basketball. One must keep track of one's four teammates, the five opponents, two baskets, the ball, the coach, and a stand full of noisy fans. The burden is on advocates of the activity-based approach to show how this task can be accomplished with minimal representations.

## How do we see what is not there?[1]

Lothar Spillmann[a] and John S. Werner[b]

[a]*Institute of Biophysics and Radiation Biology, University of Freiburg, D-79104 Freiburg, Germany;* [b]*Department of Psychology, University of Colorado, Boulder, CO 80309-0345.* **spillmann@uni-freiberg.de jwerner@clipr.colorado.edu psych-www.colorado.edu/faculty/werner.html**

**Abstract:** Pessoa et al. provide a valuable taxonomy of perceptual completion phenomena, but it is not yet clear whether these phenomena are mediated by one kind of neural mechanism or more. We suggest three possible neural mechanisms of long-range interaction to stimulate further