Perceiving Persisting Objects

Brian J. Scholl

Department of Psychology, Yale University
New Haven, CT, USA

Introduction

One of the central lessons we have learned from the emerging discipline of cognitive science is that most seemingly obvious and effortless mental tasks — e.g. recognizing faces and understanding speech — are in fact the result of incredibly complex feats of cognitive processing. Our minds not only make such accomplishments possible, but they do so in a seemingly effortless way that hides their complexity and difficulty from us. As such, much of the most important research in cognitive science begins by pointing to an aspect of our mental lives that we typically take completely for granted, and then asking how it is possible.template incorporates the following styles, which should be used where possible.

This talk will explore one such seemingly obvious aspect of our mental lives: our perception of the world in terms of persisting objects. Visual experience consists of more than individual snapshots of the world: we must also bind these discrete snapshots over time into a coherent dynamic experience. Not only must we perceive objects, but we must see them as the *same* objects over time and motion. While a tremendous amount of research has explored static object representations, surprisingly little has focused on the factors which underlie the representation of persisting objects, beyond low-level motion mechanisms.

Here I will talk about several different strands of research on the processes which underlie the perception of objects which persist through time and motion. Each of these research strands will involve perceptually salient demonstrations:

Sustained inattentional blindness

Much of the work discussed in this talk will involve the dynamics of attention. Because of the sheer amount of incoming visual information, we cannot process every object in our visual fields equally; rather, perception must be inherently *selective*, making us consciously aware of only some objects at any given moment. Such selection is often realized by the application of visual *attention*, which serves as a critical 'gateway' to conscious awareness. To emphasize the importance of attention in perceiving persisting objects, I will begin by discussing a striking *failure* of object persistence: when engaged in an attentionally demanding task, we can completely fail to see otherwise salient visual objects and events which occur right in front of our eyes — even when those objects are visible and in motion for up to 5 seconds. I will describe an demonstrate such phenomena, employing our 'sustained inattentional blindness' paradigm. To emphasize that this failure to perceive persisting objects is not merely an academic curiosity, I will also describe our recent work which has shown that the incidence of sustained inattentional blindness can increase by more than 50% when subjects are simultaneously talking on a cellular phone — a result with obvious practical significance.

Attention and object persistence: Going out of *sight* vs. going out of *existence*

When we do become consciously aware of objects, we must still bind momentary percepts into persisting object representations. In order to explore the processes which mediate object persistence, we have employed the multiple object tracking (MOT) task, wherein subjects must attentionally track a number of independently and unpredictably moving identical items in a field of identical distractors. Since the items are featurally identical during their motion, subjects can only succeed in this task by using attention to track them through the motion interval. What display features mediate the ability of subjects to track objects through time and motion in this task? To explore this question, we had subjects perform MOT in displays where the items frequently passed behind occluders. Subjects were able to successfully track even when the items were briefly (but completely) occluded at various times during their motion, suggesting that occlusion is taken into account when computing enduring perceptual objecthood. Unimpaired performance in the context of these occluders, however, required the presence of accretion and deletion cues along fixed contours at the occluding boundaries. Performance was significantly impaired when items were present on the visual field at the same times and to the same degrees as in the occlusion conditions, but disappeared and reappeared in ways which did not implicate the presence of occluding surfaces — e.g. by imploding and exploding into and out of existence, instead of accreting and deleting along a fixed contour. This pattern of results confirms that the circuits responsible for the 'attentional pursuit' of the items in this

task are not simply robust in the face of any interruption in spatiotemporal continuity, but rather have a specific tolerance for interruptions consistent with occlusion. In other words, the local dynamics of items during brief disappearances help define what 'counts' as a persisting visual object: items which disappear and reappear via accretion and deletion along a fixed contour are represented as persisting objects, and can be tracked in MOT, whereas those which disappear in other ways cannot be so tracked, as the disappearances seem to disrupt their continuing representation as the same object.

Why so slow? The role of *speed* in maintaining object persistence through occlusion

As noted above, issues of object persistence become particularly salient when persisting objects are not continuously visible—e.g. when they are passing behind occluding surfaces. For example, suppose a moving object passes behind an occluding surface, disappearing at one side of the occluder only to reappear at the opposite side a few moments later. How does the visual system decide if these events involved a single object or two distinct objects? The visual system may anticipate an object's reappearance from occlusion based on the moving object's spatiotemporal trajectory. We have explored such situations by adding various anomalies into such occluded trajectories, for example making the object appear at the right place but the wrong time, or the right time but the wrong place. Here I will describe how speed anomalies impact (and in some cases fail to impact) object persistence. Intuitively, a judgment of persisting objecthood may seem less likely when the reappearing object appears 'too slowly' (as if it had suddenly become 'stuck' behind the occluding surface) or 'too quickly' (e.g., reappearing instantly from the other side of the occluder, as if it had suddenly and coincidentally sped up while traversing the occluder). We have explored whether an object's speed is used to compute persisting objecthood in the context of the multiple-object-tracking task. Subjects tracked a subset of identical randomly moving objects about a computer screen while the objects passed behind occluders; the persisting identities of these objects had to be maintained in order to indicate the target subset at the end of the motion phase. Performance was predictably impaired by anomalous 'slow-downs' during occlusion events; the longer the anomalous delay, the worse the tracking. However, tracking was surprisingly resistant to anomalous 'speed-ups': subjects performed equally well in the 'speed-up' condition as in the baseline tracking condition, even when the occluders were many times wider than the objects themselves . These results constrain theories of the nature of the processes which compute object persistence in such dynamic situations.

Attentive tracking of objects vs. substances

In another set of studies on object persistence using the multiple object tracking task, we explored a contrast which has played a critical role in research on object-based processing in both the infant cognition and word-learning literatures: rigid cohesive objects vs. nonsolid substances. Whereas objects may move from one location to another, a nonsolid substance must *pour* from one location to another. We explored whether attentive tracking processes are sensitive to dynamic information of this type. Using MOT, we showed that subjects can easily track 4 in 8 identical unpredictably-moving entities which move as discrete objects from one location to another, but cannot track similar entities which noncohesively 'pour' from one location to another — even when the items in both conditions follow the same trajectories at the same speeds. Other conditions reveal that the inability to track multiple 'substances' stems not from the violations of rigidity or cohesiveness per se, since subjects are able to track multiple non-cohesive collections and multiple non-rigid deforming objects. Rather, the impairment is due to the dynamic extension and contraction during the 'substance-like' motion, which render 'the' location of the entity ambiguous. These results demonstrate a convergence between visual tracking processes and infant cognition, and in general help to clarify what can count as a persisting dynamic 'object' of attention.

Causal capture: Computing 'what went where'

The research described above was largely involved with the factors that mediate *whether* an object representation persists. In displays with multiple persisting objects, however, another critical issue is the nature of the visual processes which mediate the representation of 'what went where'. This is a type of visual 'correspondence problem': beyond computing *that* an object has persisted, the visual system must also decide exactly *how* it has persisted, and where it has gone. We have explored these issues in simple displays where it is clear that two objects are persisting, but it is often ambiguous as to 'which is which'. One such project has explored ambiguous percepts in dynamic displays involving *causal events*.

In addition to perceiving the colors, shapes, and motions of objects, we can also perceive higher-level properties of visual events. One such property is causation, as when you see one object cause another object to move by colliding with it. We have discovered a striking new type of contextual effect on the perception of such collision events. Consider an object (A) which moves toward a stationary object (B) until they are adjacent, at which point A stops and B starts moving along the same path. Such 'launches' are perceived in terms beyond these kinematics: as noted in Michotte's classic studies, we perceive A as being the *cause*

of B's motion. When A and B fully overlap before B's motion, however, observers often see a completely non-causal 'pass', despite salient featural differences: one object remains stationary while another passes over it. In the presence of a distinct nearby launch event, however, this stimulus is 'captured': it too is now irresistibly seen as causal. This contextual capture requires that the context event be present for only 50 ms surrounding the 'impact', but is destroyed by only 200 ms of temporal asynchrony. I will demonstrate such cases, and others, that help define the rules which the visual system uses to construct percepts of seemingly high-level properties like causation — and to compute 'which went where' in ambiguous dynamic displays.

Conclusions

These projects — and others which will also be described — each explore a different aspect of one of the most critical but underappreciated tasks of the visual system: binding momentary discrete percepts into representations of *persisting* objects. Collectively, the results of these studies begin to reveal: (1) the degree to which we can fail to consciously perceive persisting objects; (2) several of the constraints involved in determining *whether* an object is the same object as some earlier perceived object; and (3) the additional constraints involved in determining which currently visible objects should be bound with which previously viewed objects.