Chapter 5

Do the same principles constrain persisting object representations in infant cognition and adult perception?: The cases of continuity and cohesion

Erik W. Cheries, Stephen R. Mitroff, Karen Wynn, & Brian J. Scholl

5.1 Introduction

In recent years, the study of object persistence—how the mind identifies objects as the same individuals over time—has been undergoing a renaissance in (at least) two different fields of cognitive science. First, vision scientists have come to understand some of the principles that control the construction, maintenance, and destruction of object representations in mid-level vision. Second, developmental researchers have identified several principles of ‘core knowledge’ that constrain object permanence in infants. These two fields have traditionally operated largely independently, but some researchers have suggested that they may in fact be studying the same underlying mental processes.

This interesting idea has been used in the past to interpret various empirical results in each field, but the real promise of this approach lies in its ability to drive further progress by generating novel predictions that can then be tested in both fields. The hope is that this approach could spark a useful feedback loop of sorts: for example, infancy research may give rise to specific predictions for adult perception experiments, whose subsequent results may in turn give rise to additional specific predictions for infant cognition experiments. To the extent that this strategy succeeds—confirming ever more specific predictions as hypotheses are carried back and forth across these fields—we may obtain support for the idea that these two fields are studying the same thing.
In this chapter we describe two examples of attempts to implement this strategy in practice, while studying two of the most salient principles of core knowledge: continuity and cohesion. In each case, earlier infant research was used to motivate adult perception experiments, which were in turn used to generate and test more specific predictions in further infant studies. These case studies illustrate the utility of bridging the gaps between these two fields, as our knowledge of each is deepened as a result of exploring the connections. In particular, this process has revealed new ways how violations of these principles of core knowledge in turn have deleterious effects on the underlying object representations themselves. In addition, the results in each case are consistent with the possibility that representations of persisting objects in each domain are controlled by the same principles, and perhaps even the same underlying processes.

5.1.1 Object persistence

Many of the mind’s most extraordinary accomplishments are achieved with deceptively little conscious effort. A striking example of this is the apparent ease with which we perceive the many objects that populate our everyday visual experience. By merely opening our eyes, it seems that the undifferentiated wash of light captured on the surface of our retinas is instantly transformed into representations of discrete objects. However, the principal lesson from research in cognitive science (underscored by the amount of attention given to the topic in this book!) is that even a seemingly effortless ability such as object perception is in fact the result of complicated cognitive processing beneath the surface.

This is true not only for the perception of objects but also for the perception of persisting objects—our ability to perceive objects as the very same individuals over time, motion, featural changes, and despite interruptions such as occlusion (for a summary, see Scholl & Flombaum, in press). Without such processing, visual experience would be incoherent: we would be able to perceive objects, but it would be as if the world was created from scratch at every instant. In addition, recent empirical work has shown that computations of object persistence are critical for understanding many other processes such as visual memory (e.g., Cheries et al., 2006; Flombaum & Scholl, 2006), implicit learning (e.g., Fiser et al., 2007), numerical cognition (e.g., Feigenson et al., 2004), motion perception (e.g., Dawson, 1991), search and foraging (e.g., Flombaum et al., 2004; Santos, 2004), and even visual awareness in the first place (e.g., Mitroff & Scholl, 2005; Moore & Lleras, 2005).

In recent years, researchers in several areas of psychology have made great strides in elucidating the principles that guide the computation of object persistence. Here, we focus on two such domains: infant cognitive development and adult vision science.1 Collectively, the results of these studies suggest that there are mental mechanisms for computing object persistence that are primitive in at least two senses: (1) they occur relatively early in perceptual processing, such that they are largely ‘hardwired’ parts of perception that then constrain later cognition; and (2) they arise early in human development.

5.1.2 Object persistence in infant cognition

Research in infant cognition across the last two decades has generated a wealth of evidence that even young infants (a) experience the world in terms of discrete persisting objects and (b) have relatively sophisticated expectations about how objects in their physical environment should behave (for a review, see Santos & Hood, Chapter 1). For instance, infants expect that objects hidden from view continue to exist (Baillargeon, 1987; Baillargeon & DeVos, 1991; Shinskey & Munakata, 2003), move in spatially continuous paths over time (Spelke et al., 1995; Xu & Carey, 1996), occupy physically distinct locations from other objects (Baillargeon et al., 1985; Spelke et al., 1992), and maintain cohesive boundaries over time (Needham, 1999; Spelke & van de Walle, 1993). These findings are collectively thought to reflect a set of ‘core knowledge’ principles involving (among others) continuity, solidity, and cohesion; these principles guide infants’ expectations about how objects will behave and serve as the foundation of our physical understanding of the world throughout adulthood (Carey & Spelke, 1996; Spelke, 2000; Spelke & Kinzler, 2007).

A growing number of studies have also demonstrated that infants possess sophisticated numerical abilities, and these studies provide a particularly strong demonstration of infants’ ability to represent objects as persisting individuals. After all, establishing a correct numerical representation of any array requires that each of the individual objects in that array be represented as the same individual over time. As a simple demonstration of this fact, imagine a ball rolling behind a screen and emerging from the other side. How many objects were there? Adult observers, of course, would say (and see) ‘one’ (Burke, 1952; Michotte et al., 1964/1991). However, if we were unable to represent the pre- and post-occlusion encounters as the same individual, we might instead say and see ‘two’ objects. As this illustrates, the mind must somehow collate the multiple ‘snapshots’ of the occluding object before successfully representing the number involved.

1 Although we limit ourselves to these two areas of psychology in this chapter, note that the nature of object persistence has also been the focus of a great deal of recent research in other areas of cognitive science, such as computer science (for a review, see Yilmaz et al., 2006) and philosophy (for a review that highlights connections with psychological research, see Scholl, 2007).
Even at an early age, infants show precocious numerical abilities. For instance, infants can represent both small and large numbers of objects (Antell & Keating, 1983; Feigenson, 2005; Starkey & Cooper, 1980; Strauss & Curtis, 1981; van Loosbroek & Smutsman, 1990; Wynn et al., 2002; Xu, 2003; Xu & Spelke, 2000), actions (Sharon & Wynn, 1998; Wood & Spelke, 2005; Wynn, 1996), and sounds (Lipton & Spelke, 2004). Moreover, they not only represent the number of individuals under occlusion but also successfully update this representation in their minds as individuals are either added to or subtracted from the set, in a form of primitive arithmetic (Cheries & Wynn, 2007; Koecchin et al., 1997; McCrink & Wynn, 2004; Simon et al., 1995; Wynn, 1992).

### 5.1.3 Object persistence in adult visual cognition

We interpret the world in terms of persisting objects in part because that matches the inferences that we make during reasoning and other forms of higher-level cognition (e.g., Rips et al., 2006; Scholl, 2007). However, we also perceive the world in terms of persisting objects in the first place, driven by more hardwired principles of persistence that appear to be a part of visual processing. These visual processes may even cause us to perceive persisting objects even when we simultaneously judge that in fact there is no persistence—as occurs, for example, in some cases of apparent motion (Anstis, 1980; Wertheimer, 1912/1961) or the tunnel effect (Flombaum & Scholl, 2006; Michotte et al., 1964/1991).

Much of the work on object persistence in adult visual cognition occurs at the level of what is often called mid-level vision (Nakayama et al., 1995). At this level, objects may be represented as the same individuals over time despite changes to both their lower-level visual features (e.g., red, round) and their higher-level category descriptions (e.g., ‘apple’). These mid-level object representations—often referred to as object files (Kahneman & Treisman, 1984; Kahneman et al., 1992)—may store information about objects’ visual features, but their construction, maintenance, and destruction appears to be controlled by purely spatiotemporal aspects of how the objects move across space (Gao & Scholl, in press; Kahneman et al., 1992; Mitroff & Alvarez, 2007; for a review, see Flombaum et al., Chapter 6).

Some of the evidence for this system of representation comes from the object-reviewing paradigm (Kahneman et al., 1992; see Fig. 5.1).

In a typical task, adult subjects are shown a video screen with two small outlined boxes, and distinct single letters (or other distinguishable features) are briefly presented in each box. After the letters disappear, the (now empty) boxes move to different locations on the screen, after which a single probe letter appears on just one of the boxes. The subject’s task, in one variant of this paradigm (see Noles et al., 2005) is to use a key ‘press’ to indicate whether

![Fig. 5.1 Sample displays used in the object-reviewing paradigm (Kahneman et al., 1992). Subjects see two preview letters in an initial display and a single probe letter in a final display, and must simply indicate whether the final probe was present in the initial preview display. Responses are faster when the probe letter had initially appeared in that same object during the preview display, compared with when it had appeared in the other object. In the static displays, the target is seen as the same object as one of the previews, because it appears on the same object, in the same location. Objecthood and location are uncorrelated in the moving displays.](image-url)
5.1.4 Do the same principles constrain object persistence in infant cognition and adult perception?

Despite the wealth of recent research on object persistence in infant cognition and adult perception, these two fields have traditionally operated largely independently. Nevertheless, some researchers have suggested that they may be related—reflecting the same underlying principles of persistence, and perhaps even some of the same underlying processes (e.g., Carey & Xu, 2001; Chiang & Wynn, 2000; Feigenson, Carey & Hauser, 2002; Scholl & Leslie, 1999). These suggestions have been based in part on various analogies between the results that have been obtained in the two fields. Three examples (drawn from Scholl & Leslie, 1999):

Spatiotemporal priority: Both fields have converged on a core principle that guides the creation and maintenance of persisting object representations—the principle of ‘spatiotemporal priority’. When identifying objects as the same individuals over time, the visual system appears to rely on their spatiotemporal histories—that is, where, when, and how they were encountered—to a greater degree than their visual surface features (for a review, see Flombaum et al., Chapter 6).

Numerical limits: The ability to maintain persisting objects in both fields appears to be numerically constrained to only a small number (3–4) of simultaneously active representations (e.g., Feigenson et al., 2002; Pylyshyn & Storm, 1988).

Surviving occlusion: Persisting representations as studied in both fields appear to support online visual experience of the world, but they are not entirely fleeting; rather, such representations can survive periods of complete occlusion (e.g. Scholl & Pylyshyn, 1999; Spelke et al., 1995).

Exploring such analogies is valuable, insofar as it may help to theoretically bridge these two otherwise-disparate disciplines. If nothing else, doing so can reveal a wealth of new ideas for experiments in each area, as researchers discover how strikingly similar questions have been explored in various ways in each domain. However, the real promise of this (or any such) project lies in its ability to drive further concrete progress in each field. And so we might ask: can attending to the connections between these fields actually lead to a greater understanding of how object persistence is achieved in infant cognition and/or adult perception?

We think the answer to this question is ‘yes’, and in the remainder of this chapter, we discuss two case studies of such progress in practice. In each case, the initial spark for our research was drawn from studies of object persistence in infant cognition. These studies prompted us to ask similar questions about how such manipulations might or might not operate in adult perception. In turn, though, the results we obtained in the adult perception experiments revealed new nuances to such processing that then deserved and received further testing in infant cognition—implementing a type of feedback as we moved back and forth between these domains. The results we obtained along each step of this process deepened our understanding of some of the core principles of persistence in a way that would have been difficult to discover without bridging these disciplines. And the strikingly similar patterns of results that we observed across these areas is consistent with the possibility that we have in fact been studying the same underlying principles and/or processes.

5.2 Case study 1: Continuity through occlusion

A basic fact about the world is that objects cannot simply go in and out of existence over time: for two objects encountered at different locations to be subsequent stages of the same individual, there must be a spatiotemporally continuous path between them. If an object disappears at one location, and an object immediately appears at a different spatially separated location, then those two instances cannot be the same object, because physical laws do not allow for that sort of thing (at least at the spatial and temporal scales that characterize our everyday interaction with objects). As such, tracking an object’s continuous movement across space could serve as a reliable guide for perceiving it as the same persisting individual over time. One problem with this constraint in practice, however, is that our eyes frequently lose perceptual contact with many of the objects we may be tracking. Along with the effects of blinking, perhaps the most severe form of this type of interruption is occlusion, where one object may disappear entirely behind another object before reappearing. In order to perceive objects as they really are—individuals that continue to persist over time—rather than how they literally appear to our eyes, our mind must account for ecologically valid disappearances such as occlusion.

5.2.1 Initial infant research

There is now ample evidence that the ability to maintain object representations throughout visual interruptions such as occlusion develops quite early in ontogeny (for reviews, see Baillargeon, 1999; Santos & Hood, Chapter 1). For instance, even newborn infants are able to perceive partly occluded objects...
as complete wholes rather than as the fragmented parts that are immediately perceivable (Valenza et al., 2006; cf., Amso & Johnson, Chapter 9), and by at least 3.5 months, infants expect objects that have completely disappeared to persist and move continuously during momentary disappearances (Aguirre & Baillargeon, 1999, 2002; Baillargeon & DeVos, 1991; Spelke et al., 1995). In one study, for example, the looking-time behavior of 5-month-old infants who observed a rod that moved sequentially behind two screens (see Fig. 5.2a) indicates that they perceive the event in terms of a single persisting object. In contrast, however, they instead seem to perceive two distinct objects if the rod does not traverse the space between the two screens (Fig. 5.2b; Spelke et al., 1995). This simultaneously illustrates (a) the operation of a constraint on persistence based on spatiotemporal continuity and (b) the ability of infants to keep track of objects through periods of occlusion.

5.2.2 Step 1: Testing for persistence through occlusion in adult perception

Directly inspired by such demonstrations in the infant cognition literature (and especially by Spelke et al., 1995), one of our early experiments (Scholl & Pylyshyn, 1999) sought to determine whether object tracking in adult perception is also able to survive periods of occlusion—or whether constant perceptual contact must be maintained with objects in order to track them over time as the same individuals. In an MOT task, adult subjects had to track 4 out of 8 moving objects as they moved about a display that also contained two occluders. Each time a moving object intersected an occluder, it disappeared, and later reappeared from the opposite side of the occluder. Across trial types, the specific manner of these disappearances and reappearances was varied (see Fig. 5.3). On Occlusion/Disocclusion trials, whenever an object intersected the edge of an occluder, it would disappear gradually at its leading edge (as if it were moving behind the occluder) and then later gradually reappear along its trailing edge (see Fig. 5.3a). Tracking performance on such trials was excellent, and not significantly different from the performance obtained when the objects simply moved atop the ‘occluder’ contours, so that they were always visible. Thus, it seems as if persisting representations in adult perception, similar to those in infant cognition, can readily survive occlusion.²

² Later experiments (e.g., Flombaum et al., 2008; Scholl & Feigenson, 2004) have sometimes found a slight impairment for tracking through occlusion compared with tracking without any occluders—but in all cases, the magnitude of this impairment has been small, and tracking performance in both conditions has been excellent.
The ability to track well in Occlusion/Disocclusion trials does not itself provide evidence for any principles that control object representations; rather, such success could simply reflect a general tolerance of tracking through any form of brief interruption. These adult perception experiments also tested other types of manipulations, however, beyond the familiar types of occlusion. On Implosion/Explosion trials, whenever an object intersected the edge of an occcluder, it would shrink to a central point at the entering edge of the occcluder, and then later expand out of a central point at the exiting edge (see Fig. 5.3b).

These trials were constructed so that the objects were always visible and invisible for the same amounts of time as in the Occlusion/Disocclusion trials, and the disappearances always occurred at the same times and at the same rates. Nevertheless, performance on Implosion/Explosion trials was radically impaired relative to Occlusion/Disocclusion trials, and was often near chance. (Moreover, similar results occur with Implosion vs. Occlusion in experiments involving the tunnel effect that do not require attentive tracking or explicit responses involving persistence over time; Flombaum & Scholl, 2006.)

In sum, when subtle visual cues indicated that the momentary disappearances reflected the objects going out of sight, the disappearances did not affect tracking—but when these cues indicated that the objects were going out of existence, subjects were no longer able to track them as persisting individuals, despite their brute visual similarity to the ‘trackable’ displays (Scholl & Feigenson, 2004; Scholl & Pylyshyn, 1999). Critically, notice that this inability did not reflect subjects’ beliefs or strategies: they knew perfectly well what was going on during the Implosion/Explosion trials, and would have preferred to ignore those cues. But they could not—indicating that this particular type of principle of spatiotemporal continuity is encapsulated from our beliefs and preferences, and is part of the basic processes that help to generate visual experience.

5.2.3 Step 2: Testing occlusion versus implosion in infant cognition

The pattern of results obtained in the studies described in the previous section succeeded in taking a prediction from the infant cognition literature and testing (and confirming) it in the adult perception literature. In our attempt to unconfound the specific principle of continuity through occlusion from a more general tolerance for interruptions, however, we also uncovered a new effect of the manner of disappearance at occluding boundaries. This effect (operationalized as occlusion vs. implosion) may also reflect a type of principle governing the maintenance of persisting object representations: only objects that disappear in specific and ecologically valid ways can be maintained through interruptions. If this type of principle is truly primitive in the mind, however—and if the adult perception and infant cognition literatures are exploring related processes—then we should similarly find that infants’ representational abilities are constrained in this fashion.

In fact, two recent studies have demonstrated that such cues can influence infants’ representations of hidden objects. First, infants’ ability to predict whether a ball will reappear after disappearing behind an opaque screen depends upon such cues (Bertenthal et al., 2007): infants who observe a ball that disappears via ‘occlusion’ (as in Fig. 5.3a) make significantly more saccades to the opposite side of the screen than infants who see the ball ‘implode’ behind the screen (as in Fig. 5.3b). In this way, the cue that signals an object’s spatiotemporal continuity controls whether infants anticipate that the object will reemerge from the other side of a screen. Similarly, an electroencephalography (EEG) study with 7-month-old infants has shown that such cues control activity levels in brain regions that correlate with the maintenance of object representations (Kaufman et al., 2005). Significantly more gamma-band activity was detected over these brain areas when infant subjects viewed an object being covered by a screen in a way that was consistent with occlusion, compared with trials when the object gradually dissolved as it was covered.

Both of these studies underscore how object persistence in infancy is constrained by the same visual cues that support object tracking in adult perception. However, we might still want to assess whether this manipulation disrupts infants’ representations of persisting objects per se. In particular, note that adults’ failure to track multiple imploding and exploding objects in the MOT task may reflect the disruption of two separate computations. First, the implosion cue at the entering edge of the occluder may have disrupted the visual system’s ability to infer the continued existence of the object behind the barrier (as demonstrated in the aforementioned infant studies). Additionally, however, this cue was also sufficient to destroy adults’ ability to track the emerging object as the same individual once it reappeared. This second result is not a necessary consequence of the first, because the continuity of speed, trajectory, and featural similarity could in principle be sufficient for the numerical identity of the object to be maintained once it reappears. ('I thought it was gone, but wait—it’s back now!')

Does this second type of disruption—the inability to reacquire an object as the same individual once it does reappear, after an Implosion/Explosion event—also occur in infant cognition? To test this, we adapted the adult
MOT task for infant subjects (Cheries et al., in preparation; see also Cheries, Feigenson et al., 2008). Ten-month-old infants were habituated to dynamic displays of either 2 or 3 randomly moving identical items, which disappeared and reappeared from behind two long and narrow occluders. What varied across experimental conditions (and subjects) was precisely how these objects disappeared behind the opaque barriers: they either occluded and disoccluded (Fig. 5.3a) or imploded and exploded (Fig. 5.3b) as in the previous adult perception studies. The critical question was whether infants would be able to establish representations of a stable number of individuals across these habituation trials. To assess this, infants in both types of disappearance conditions were shown the identical displays of 2 or 3 moving objects at test but without any barriers or disappearances. If infants maintained the number of individuals across habituations trials, then infants who were habituated to 2 objects should respond by looking longer at test displays of 3 objects and vice versa.

As predicted, infants who were habituated to objects that disappeared behind the barriers via ‘occlusion’ and ‘disocclusion’ successfully dishabituated to the test displays containing a novel number of objects (see Fig. 5.4a). In contrast, infants who were habituated to object arrays that disappeared and reappeared from behind the screens via ‘implosion’ and ‘explosion’ failed to differentiate the number of objects present in the proceeding test trials (see Fig. 5.4b). Apparently, infants were unable to perceive a stable number of individuals across habituation in this condition, even though the objects moved and disappeared equally gradually and at the same rates as the objects in the occlusion condition.

These results are consistent with the interpretation that occlusion is a cue that an object has gone out of sight, whereas implosion is a cue that an object has gone out of existence. However, an additional possibility is that the imploding motion pattern is so anomalous and distracting that infants were unable to successfully represent anything at all about the display over the course of habitation, due to a general distraction. After all, infants have no experience observing objects behaving in this manner.

We tested these two possibilities by determining whether infants are able to represent a nonnumerical property of the objects in the ‘implosion/explosion’ condition. Infants were again habituated to displays of moving objects that imploded and exploded from behind the barriers. Some infants were shown 2 or 3 disc-shaped objects as before, whereas another group of infants were habituated to 2 or 3 square-shaped objects. After this habituation phase, infants saw test trials that always contained the same number of objects they were habituated to, but the shapes of the object were alternated across test trials.

This design allowed for a simple test of whether infants were able to encode and represent the shape of the objects during habituation, despite the potentially distracting implosion and explosion cues. Indeed, infants who were habituated to imploding/exploding discs were found to successfully dishabituate to test displays containing the same number of squares and vice versa. This rules out the possibility that infants are just generally distracted by the implosion/explosion motion, such that they fail to represent any feature of the objects from the habituation trials at all. Rather, the disruptive effect of the anomalous disappearance cue is specific to maintaining representations of the objects’ numerical identity over time.

The experiments reviewed in this section illustrate the utility of bridging the adult perception and infant cognition literatures in the context of object persistence in general, and spatiotemporal continuity in particular. As a result, we have learned about the role of spatiotemporal continuity in adult perception,
and we have learned about the importance of the manner of disappearances and reappearances in infant cognition. In particular, the most recent step in this feedback loop has revealed that violations of spatiotemporal continuity via implosion and explosion cues serve not only to guide orienting (Bertenthal et al., 2007) but also to actively disrupt the underlying numerical representations of infants during passive viewing (Cheries, Feigenson et al., 2008, in preparation).

5.3 Case study 2: Cohesion violations

One of the most powerful principles determining whether something can be tracked as the same persisting individual over time may be that of cohesion: objects must maintain rigid boundaries and internal connectedness as they move over time. This may also be the most intuitive defining characteristic of what it means to be an object in the first place. For instance, to determine whether something is an object, you might just ‘grab some portion of stuff and pull: all the stuff that comes with you belongs to the same object; the stuff that remains behind does not’ (Bloom, 2000, p. 94; Pinker, 1997). In this way, a cohesion violation differs from other core principles (e.g., continuity) in that an entity's noncohesive motion may both (a) violate how an object is expected to move and, in some conditions, (b) call into question whether the entity was correctly categorized as an object in the first place.

5.3.1 Initial infant research

One way that the principle of cohesion may disrupt our persisting representations is by causing some entity to be categorized as a nonsolid substance rather than an object. The 'core knowledge' framework (Spelke, 2000; Spelke & Kinzler, 2007) stipulates that principles such as solidity, continuity, and cohesion selectively apply only to the domain of physical objects per se—and as such, this framework predicts that infants will have few (if any) expectations about events involving entities that are categorized as noncohesive substances. Indeed, infants demonstrate a surprising lack of expectation about physical events involving substances, compared to their apparent savviness when it comes to objects. Even at the same age when infants can successfully remember the precise number of objects that have disappeared behind an occluder (e.g., Wynn, 1992), they fail to demonstrate any expectation about whether a poured substance, such as sand, should continue to exist (Huntley-Fenner et al., 2002; for reviews, see Rosenberg & Carey, Chapter 7; vanMarle, 2004). In this way, the work on noncohesive substances provides some of the strongest evidence to date of the domain-specific nature of our early physical reasoning abilities.

Another way that the principle of cohesion might constrain our representations is by guiding our expectations about how objects ought to behave. Nearly all of our visual experience supports the notion that objects maintain their boundaries and move as connected wholes over time. In fact, even young infants expect that objects will remain cohesive over time. For example, in a display wherein a hand grabs an object from above, infants look longer at events where an object appears split apart due to the grabbing, compared with similar grabbing events involving two objects that already appear separated (Needham, 1999; Spelke & van de Walle, 1993; Xu et al., 1999). This demonstrates one way in which infants' expectations about objects are constrained by cohesion: this principle requires that a solid-looking object will maintain a single cohesive boundary over time.

Previous work from our laboratories has also demonstrated how certain complex types of cohesion violations impair infants' expectations about object persistence. For example, infants' fail to detect the magical disappearance of a pyramid object that has been disassembled into 5 separate blocks, reassembled into its pyramid shape, and then pushed behind a screen (Chiang & Wynn, 2000). This, and a series of control studies, showed how infants' persisting object representations are perturbed by the process of reconstituting a single object as a collection of multiple objects. This could be due to the type of categorization effects discussed earlier, where an object breaking into five smaller pieces is recategorized as a substance that is then no longer bound by the principle of continuity. Alternatively, this disruption may have been due to the fact that a single object file was attempted to be reassigned to multiple objects that exceed the capacity limits of attention in infants (for a review, see Zosh & Feigenson, Chapter 2).

5.3.2 Step 1: Testing for cohesion violations in adult perception

In adult perception research, objects are often contrasted with spatial regions or visual surface features (for a review, see Scholl, 2001), but they had never to our knowledge been contrasted with nonsolid substances—as was so prominent in the infant cognition research noted in the previous section. This observation by itself illustrates the usefulness of exploring object persistence across subfields of cognitive science—because in this case an entire class of entities had been ignored in one field but was prominent in another.

Directly inspired by demonstrations of the importance of cohesion in the infant cognition literature, we sought to determine whether this principle also constrains persisting object representations in adult perception. In an initial study, we used the MOT paradigm (described in Section 5.1.3) to test
adults' ability to track objects that move from location to location in a non-cohesive manner (vanMarle & Scholl, 2003). As in the standard MOT task (e.g., Pylyshyn & Storm, 1988; Scholl & Pylyshyn, 1999), adults were presented with a homogeneous array of 8 rectangular objects and were asked to track a particular subset of 4. On Object trials, the shapes moved as in typical MOT experiments, maintaining their boundaries over time while their locations shifted. On Substance trials, however, the objects moved by disintegrating into streams of pixels, as if being poured from one location to the next. Although this 'pouring' motion traced exactly the same trajectories and shifted at the same rates as various other control conditions, subjects' tracking performance was significantly impaired.

This MOT research suggested that tracking (nonsolid) substances was markedly more difficult than tracking discrete bounded objects, but the underlying reason for this difference was unclear. Rather than being due to some sort of categorization as a substance, the 'substance tracking' impairment may have been due to a characteristic dynamic property of nonsolid substances (vs. objects), which is that they do not have a consistent center of mass over time, to which attention might be locked during tracking (vanMarle & Scholl, 2003; for evidence that tracking does prioritize such a location, see Alvarez & Scholl, 2005). As a result, however, it was an open question whether cohesion violations directly impaired persisting object representations per se.

To study this more specific question, we employed the object-reviewing paradigm as discussed earlier (see Section 5.1.3), and did so with an extremely simplified and ‘pure’ form of cohesion violation: a single object splitting smoothly into two. This seems like a rather pedestrian type of event, but note that it introduces a critical type of ambiguity. Suppose that you witnessed such an event and were then asked where the initial object was. None of the obvious possible answers (it’s just one of the resulting ‘twins’; it’s neither of them; it’s both of them) seems especially satisfying.

To explore the visual system’s answer to such questions, adult subjects viewed two circular objects that moved from left to right across a display. During the motion, one of the objects 'split'; as it continued to move, it gradually separated into two identical resulting objects (Mitroff et al., 2004; see Fig. 5.5). Critically, this splitting was smooth and symmetrical, so that there was nothing to bias one of the resulting objects over the other. (A second nonsplitting object is also present in each display as a control.) The theoretical question then concerned what would happen to the underlying object representation of the initial object—its 'object file'—as a result of the fission. There are several possibilities that correspond to the intuitive possibilities noted in the preceding: (1) the cohesion violation might obliterate the object file’s contents entirely;

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Concurrent trials} & \text{Incongruent trials} \\
\hline
\text{Preview display} & \text{Linking motion} & \text{Target display} & \text{Preview display} & \text{Linking motion} & \text{Target display} \\
\hline
\text{Upward split} & \text{X} & \text{X} & \text{X} & \text{X} & \text{X} & \text{X} \\
\text{Downward split} & \text{X} & \text{X} & \text{X} & \text{X} & \text{X} & \text{X} \\
\text{Straight motion} & \text{X} & \text{X} & \text{X} & \text{X} & \text{X} & \text{X} \\
\hline
\end{array}
\]

\textit{Fig. 5.5} A subset of the trial types from a study of cohesion violations using the object-reviewing paradigm (Mitroff et al., 2004). In each case, subjects simply responded as quickly as possible whether the final letter appeared anywhere in the initial preview display on that trial. Objects either traveled a straight trajectory, a curved trajectory, or smoothly split into two separate objects. (The actual experiment also included No-Match trials for each condition, in which the final probe letter was neither of the initially presented letters, and factors such as the relative orientations were always counterbalanced and randomized, such that the final probe letter could appear on any final object.) As described in the text, these cohesion violations produced considerable costs in visual processing, relative to controls that are not depicted here.

(2) the information about the preview letter might survive the 'splitting' intact, but stay bound to only one of the two resulting objects, indicating that object files cannot themselves split into two; or (3) the object file's contents might essentially be 'copied' to both of the two resulting objects.

In fact, this cohesion violation detrimentally affected the subjects’ ability to maintain a persisting representation over time and motion in two ways. First, the magnitude of the OSPB for a 'split' object was significantly smaller than that obtained during control trials where there were similar types of motions but no splitting. Second, the OSPB for an additional third object in the display that simply moved across the screen unperturbed was completely eliminated by the existence of splitting elsewhere in the display. These costs suggest that the visual system ends up effectively splitting the actual object representations that underlie the object(s) in this event, such that one object file is eventually replaced with two object files, into which the initial contents of the first file are copied. (There is no indication in these studies of bimodal response patterns, which would suggest that the object file 'went with' only one of the post-split objects.) The fact that this process resulted in substantial costs to
visual processing, however, indicates that this event is seen as exceptional, and requires more resources to handle than is the norm in the perception of simple events.

This result demonstrated for the first time how the active maintenance of adults’ object representations can be constrained by the principle of cohesion. We may be able to readily see objects that split in half (as we commonly do for various types of food, pieces of paper, etc.), but this work suggests that the cohesion violations that accompany such events nevertheless impair forms of underlying visual processing.

5.3.3 Step 2: Testing for impairments due to simple splitting events in infant cognition

The adult perception work described in the previous section was motivated by earlier infant research, but the manipulation that we employed in the object-reviewing study—a single object smoothly splitting into two—was markedly different (and simpler) than the manipulations used in the previous infancy experiments (e.g., pouring sand or disassembled blocks). If cohesion truly operates in infancy as a principle, however, then such a simple violation should nevertheless also impair infants’ object representations. (Indeed, although such an event is so simple, that simplicity also serves to make the violation itself that much more ‘flagrant’.)

So, directly motivated by our object-reviewing experiments, we tested this prediction using a forced-choice crawling procedure that has been previously used as a direct behavioral measure of infants’ quantification abilities (Feigenson et al., 2002). In our version of this task (Cheries, Mitroff et al., 2008), 11-month-old infants witnessed a single trial of an experimenter placing a different number of graham crackers into two spatially separated containers. All infants observed the experimenter placing one graham cracker into one bucket and two graham crackers sequentially placed into the other (a 1 vs 2 cracker choice). What we varied across subjects was whether the two crackers were initially presented (a) as already-separated entities or (b) created from the experimenter splitting a larger cracker in two (see Fig. 5.6). After observing one of these events, infants were allowed to crawl toward and choose either of the containers.

We replicated previous results demonstrating that a significant majority of infants will reliably crawl to the bucket containing two crackers (Feigenson et al., 2002). However, despite controlling for the overall presentation time and hand motions across conditions, we found that infants in the ‘Split Cracker’ condition chose randomly between the two buckets (see Fig. 5.7). This result was surprising, seeing as how the crackers entered the buckets in exactly the same manner across both conditions; what varied was only how the crackers first appeared to the infant. Also, by this age, infants have had many experiences seeing crackers and other food objects split in half. Although it is possible that some infants, over time, have developed a preference for ‘whole’ crackers over broken ones, this preference was shown to have no effect in other control conditions.

Even a cohesion violation as simple as a single object splitting in half was sufficient to disrupt infants’ ability to choose the container with the greater cracker quantity. However, there are at least two different ways to characterize the nature of this disruption. Attending to a cohesion violation may impair infants’ ability to maintain representations of an object’s persisting numerical identity, or it may interfere with their ability to represent and use a
particular feature of those objects, such as their overall volumes. The distinction between these two interpretations remains ambiguous in the current study because previous research has demonstrated how infants’ choosing in this paradigm is based on the relative amount of cracker material (i.e., collapsed across individual crackers) in either of the buckets. The manual-search paradigm (Feigenson & Carey, 2003, 2005; van de Walle et al., 2000; for a review, see Zosh & Feigenson, Chapter 2), on the other hand, has been shown to bias infants’ judgments away from their expectations about surface area. Instead, this paradigm elicits infant responses based on the number of individuals that are retrieved from the box (regardless of their size) that are matched to infants’ representations of those objects on the basis of one-to-one correspondence (Feigenson & Carey, 2003).

In order to provide convergent evidence for our crawling study and better characterize the nature of the observed disruption, we have recently run an analogous study in slightly older infants using a manual-search paradigm. This allowed us to test, within subjects, whether infants’ specific numerical expectations about objects placed into a box were affected by a simple cohesion violation. In this version of the task (Cheries & Carey, in preparation), 12.5-month-old infants witnessed an experimenter place 1 or 2 Lego pieces (4 trials each) into a box that they could reach but could not see into (see Fig. 5.8). In both trial types, infants were allowed to retrieve a single Lego piece by reaching through the fabric on the front of the box. The experimenter would then code the number and duration of reaches during a 10-second coding window. If infants successfully remembered how many Lego pieces entered the box, they should reach relatively less on 1-object trials, when the box should be empty, than on 2-object trials, when there should still be one Lego piece remaining. What varied across the two blocks of trials (in a within-subjects manipulation) was whether the Lego pieces in 2-object trials were initially presented as one larger Lego piece (Fig. 5.8b) or were already separated into two (Fig. 5.8a). As in the previous study, the timing and hand motions were completely controlled across ‘Split’ and ‘No-Split’ conditions.

Replicating previous studies using this method, infants successfully reached more on 2-object trials than on 1-object trials when the two Lego pieces were initially presented as two separate objects. However, infants failed to reach significantly more on 2-object trials when the two Lego pieces were created from a larger one splitting in half, even though every other aspect of the presentation was identical. Thus, even the simple and extremely familiar act of a Lego block splitting into two was sufficient to disrupt infants’ representations of how many objects were involved in the event. These data demonstrate how infants’ persisting object representations appear to be powerfully constrained by the principle of cohesion.

The experiments reviewed in this section illustrate the utility of bridging the adult perception and infant cognition literatures in the context of object persistence in general, and cohesion violations in particular. As a result, we have demonstrated for the first time that cohesion plays a role in driving object persistence in adult perception, and that at least some similar types of cohesion violations produce impairments in the contexts of both adult perception and infant cognition. In particular, the most recent step in this feedback loop has revealed that even the simplest possible cohesion violation—of a single familiar object splitting into two—dramatically impairs the ability of infants to maintain persisting representations.

5.4 Concluding Discussion

Explaining the mind’s ability to represent objects as persisting individuals has been a major preoccupation of several fields within the cognitive sciences. The promise of the approach presented here is that empirical results from disparate subject populations and research methodologies can be compared with one another to generate novel and testable hypotheses that are unlikely to have
arisen otherwise. The goal of the current chapter was to describe a successful implementation of this approach across the historically independent fields of infant cognition and adult vision science.

In particular, this chapter has focused on how our understanding of two core principles of persistence—continuity and cohesion—has been significantly enhanced by the seesaw relationship between studies in these two fields. The initial explorations of both of these principles in infant cognition directly inspired research in adult vision science, which in turn sparked further and more specific explorations of the operation of these principles back in studies with infant subjects. This feedback loop has advanced our understanding of how these principles influence our persisting object representations in several ways.

First, without noticing the emphasis placed on such cues in the infant literature, vision scientists might not have ever thought to test for the effects that cues to occlusion or cohesion have on adults’ persisting object representations. (Indeed, as noted earlier, entire distinctions—such as objects vs. substances—have been prominent in one field but entirely absent in the other.) As a result, we have learned how these principles continue to govern the online maintenance of adults’ object representations in mid-level vision in a way that appears to be automatic and remarkably encapsulated from the observer’s high-order beliefs. Observing the operation of these principles at this level in adults may also help direct the characterization of infants’ performance along the implicit/explicit divide. Before a certain age, infants may only ‘know’ that objects should remain cohesive in the same way that adults’ mid-level visual representations ‘know’ about cohesion.

In turn, without the corresponding adult vision studies, the infant literature might not have predicted the relatively subtle cues that articulate these constraints. For example, success and failure on the adult tasks was determined by subtle differences (a) between an object gradually disappearing along a fixed contour versus disappearing just as gradually but along all contours at once or (b) between a single object splitting in half versus remaining whole. These manipulations might have first appeared as unlikely candidates to affect infants’ representations in the typical free-viewing tasks. (Indeed, one author of this chapter thought it not worth running our initial infant ‘graham cracker splitting’ study, because he found it so intuitively unlikely that it would succeed!) However, emboldened by the findings with adults, we have discovered how even the relatively subtle object behaviors that define these principles hinder object processing.

Observing the relative strength of these cues across both subject populations has also demonstrated the relative severity of the cost for violating these principles. Each of the subtle manipulations employed in these studies was found to significantly impair or destroy subjects’ ability to represent the objects involved in the event. It seems that objects that violate these principles frustrate the computations that support our ability to represent them as the same individuals over time. The representational costs observed in infant subjects were directly predicted by adults’ performance using analogous manipulations. Moreover, these deleterious effects constitute a somewhat new amendment to the traditional characterization of the operation of object constraints in infancy. Infants’ performance across a large number of physical reasoning tasks is consistent with the idea that the principles of cohesion and continuity constrain their expectations and interpretations of physical events.

For instance, infants often look reliably longer at physical events that violate these principles (e.g., Baillargeon et al., 1985), and tend to favor event interpretations that avoid such violations altogether, inferring the presence of additional objects if necessary (e.g., Spelke et al., 1995; Xu & Carey, 1996). However, these studies did not indicate whether these core principles similarly constrain the maintenance of infants’ persisting object representations per se. It is well established that infants will react by looking longer to various types of object violations, but most studies stop short of asking whether the object representations themselves incur any type of penalty. However, the studies discussed in this chapter, motivated by the adult vision results, show how violating these principles imposes additional processing demands that severely disrupt infants’ representations of the objects involved in the event. A failure to represent objects through such violations is suggestive that the computations involved in object persistence are constrained by these principles.

Finally, the studies reviewed in this chapter have also demonstrated how violating the principle of cohesion and spatiotemporal continuity produce deleterious effects that are remarkably continuous across development. It is unlikely that the subtle processing costs for such violations in adult perception are due to a lifetime of experience viewing objects as they normally behave in the world, because these effects seem to operate even more powerfully in infants who lack as much experience. This pattern suggests how these constraints may constitute the foundation of our object understanding from very early on or possibly even from birth. The effects may become less severe across development as we gradually develop auxiliary resources that can compensate for these representational limitations, but these results reinforce the ‘core’ nature of these principles of core knowledge. In summary, the two case studies presented in this chapter—focusing on the principles of spatiotemporal continuity and cohesion, respectively—highlight the benefits of explicitly and directly exploring how infant cognition research can inform adult perception research, and vice versa.
References


Scholl, B. J. (in press). What have we learned about attention from multiple object tracking (and vice versa)? In D. Dedrick, & L. Trick (Eds.), Computation, Cognition, and Pylyshyn, Cambridge, MA: MIT Press.


Chapter 6

Spatiotemporal priority as a fundamental principle of object persistence

Jonathan I. Flombaum, Brian J. Scholl, & Laurie R. Santos

6.1 Introduction

The impoverished and rapidly changing stimulation on the retina looks very different from the stable world of discrete persisting objects that populate our visual experience. To get from the features on the retina to the objects that we experience, the visual system must solve several correspondence problems. One of these problems has to do with sameness: the visual system must decide whether each bit of stimulation reflects an object that has already been encountered (which might occasion the updating of an existing object representation) or a new one (which might occasion the creation of a new object representation). This problem of object persistence has been studied with a wide array of visual phenomena and paradigms, and in several disciplines in cognitive science—including vision science, developmental psychology, and comparative cognition.

The study of object persistence in these different fields has progressed largely independently. Yet strikingly, they have converged on a core principle that guides the creation and maintenance of persisting object representations: the principle of spatiotemporal priority. When identifying objects as the same individuals over the time, the visual system appears to rely on their spatiotemporal histories—that is, where, when, and how they were encountered—to a greater degree than their visual surface features. In this chapter, we review the many contexts in which spatiotemporal priority drives computations of object persistence, and we propose explanations at several levels for why spatiotemporal priority plays this dominant role.