COGNITION

Cognition 80 (2001) 179–213

www.elsevier.com/locate/cognit

Infants' knowledge of objects: beyond object files and object tracking

Susan Carey[a,*], Fei Xu[b]

[a]Department of Psychology, New York University, 6 Washington Place, Rm 550, New York, NY 10003, USA

[b]125 NI, Department of Psychology, Northeastern University, Boston, MA 02115, USA

Received 23 February 2000; accepted 17 November 2000

Abstract

Two independent research communities have produced large bodies of data concerning object representations: the community concerned with the infant's object concept and the community concerned with adult object-based attention. We marshal evidence in support of the hypothesis that both communities have been studying the same natural kind. The discovery that the object representations of young infants are the same as the object files of mid-level visual cognition has implications for both fields. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Infants' knowledge of objects; Object files; Object tracking

1. Object individuation and numerical identity

Sensory input is continuous. The array of light on the retina, even processed up to the level of Marr's 2 1/2 D sketch (Marr, 1982), is not segregated into individual objects. Yet distinct individuals are provided by visual cognition as input to many other perceptual and cognitive processes. It is individuals we categorize into kinds; it is individuals we reach for; it is individuals we enumerate; it is individuals among which we represent spatial relations such as "behind" and "inside"; and it is individuals that enter into causal interactions and events. Because of the psychological importance of object individuation, the twin problems of how the visual system

* Corresponding author. Fax: +1-212-9954018.
E-mail addresses: scarey@psych.nyu.edu (S. Carey), fxu@neu.edu (F. Xu).

PII: S0010-0277(00)00154-2

establishes representations of individuals from the continuous input it receives and the development of these processes in infancy have engaged psychologists for almost a century.

Human language, and cognition more generally, makes a principled distinction between *individuals* and their *properties*. One of the quantificational functions of noun phrases is to denote individuals and sets of individuals, whereas predicates denote properties of those individuals. Accordingly, the literatures of metaphysics and philosophy of language distinguish between *sortals* (concepts that provide criteria for individuation and numerical identity) and *non-sortals* (Gupta, 1980; Hirsch, 1982; Macnamara, 1986; Wiggins, 1980). Similarly, the object-based attention literature (see papers in this volume) argues for a principled distinction between processes that index individuals and track them through time and processes that bind representations of features to representations of those individuals.

The study of object representations in infancy has an intellectual history independent of the object-based attention literature. Piaget's pioneering studies of object permanence were motivated by Kantian considerations of the origins of ontological commitments (*space*, *time*, *object*, *causality*). Piaget (1954), like Quine (1960), wondered how infants, assumed to be endowed initially only with sensorimotor representations, could construct representations of individual objects which exist independent of them. Notice that the issue of Piagetian object permanence is at the heart of the problem of numerical identity of objects with which one has lost perceptual contact. When we credit infants with an appreciation of object permanence, we assume that they know it is the *same object* that they saw disappear under the cloth that they are now retrieving. As is well known, Piaget believed that infants did not acquire true object permanence until 18–24 months, the end of what he called the period of sensorimotor intelligence. Even successful retrieval of objects hidden under and behind barriers at around 9 months is consistent with mere empirical rules that lead the child to predict that if an object is seen disappearing behind the barrier, an object will be found there (with no commitment as to whether it is the same object or a different one). However, there is now ample evidence, some of which we will review here, that infants as young as 2.5 months establish representations of individuated objects and track them through time, even when occluded.

Thus, both literatures, that on mid-level object-based attention and that on object representations in infancy, involve parallel problems, in particular those of the bases of object individuation and numerical identity. Recently, many have suggested that both communities have actually been studying the same psychological mechanisms; that is, that the object representations of young infants are identical to those that are served up by mid-level object-based attention (Leslie, Xu, Tremoulet, & Scholl, 1998; Scholl & Leslie, 1999; Simon, 1997; Uller, Carey, Huntley-Fenner, & Klatt, 1999). We endorse this proposal, with an important emphasis on *young*. Our paper has three main goals. First, we wish to introduce the literature on infant object representations to researchers studying object-based attention. Next, we summarize the considerations in favor of the hypothesis that the representations of the mid-level object tracking system are those that subserve object representations of young infants. Finally, we consider what practitioners of each discipline have to learn from those of the other if we accept this hypothesis. Although many of the arguments in this paper are highly speculative, we believe that this exercise will inform both communities and open new venues of empirical research.

2. Two distinct representational systems in the service of object individuation

In adults, there is *prima facie* evidence that *at least* two distinct representational systems underlie object individuation. The first is the mid-level vision system (mid-level because it falls between low level sensory processing and high level placement into kind categories) that establishes object file representations, and that indexes attended objects and tracks them through time (see the papers in this volume). This first system (called in this paper the mid-level object file system) privileges spatiotemporal information in the service of individuation and numerical identity. Individual objects are coherent, spatially separate and separately movable, spatiotemporally continuous entities.[1] Features such as color, shape, and texture may be bound in the representations of already individuated objects; they play a secondary role in decisions about numerical identity, when spatiotemporal evidence is neutral. Furthermore, a small number of attended objects may be indexed in parallel, the indexed individuals tracked through time and occlusion, the spatial relations among indexed individuals represented. Pylyshyn (2001) dubs these indexes FINSTs (FINgers of INSTantiation), for they serve a deictic function, like a finger point at an individual object. Here we adopt the assumption made by Kahneman, Treisman, and Gibbs (1992), and endorsed by Pylyshyn, about the relation between the indexing processes (Pylyshyn's FINSTs) and object files. Object files are symbols for individuals and FINSTs are the initial spatiotemporal addresses of those individuals. FINSTs might be thought of as the initial phase of an object file, before any features have been bound to it.

The second system (called in this paper the kind-based object individuation system) is fully conceptual, drawing on kind information for decisions about individuation and numerical identity. For adults, individuation is based on kind information when no relevant spatiotemporal evidence is available, as when we decide that the cup on the windowsill is the same one we left there yesterday, but the cat on the windowsill is not the same individual as the cup we left there yesterday. Sometimes kind information overrides spatiotemporal continuity, as when we decide that a person ceases to exist when she dies, in spite of the spatiotemporal continuity of her body. Property/featural changes are relevant to individuation at the conceptual level, but not on their own. Our inferences concerning the relevance of property changes to individuation are kind-relative. For example, a puppy may be the same individual as a large dog a month later, but a small cup will not be the same

[1] The exact characterization of the individuals that are indexed by FINSTs and are represented by object files is a matter awaiting empirical investigation. See Scholl, Pylyshyn, and Feldman (2001) for a first investigation into what individuals can be tracked in multiple object tracking (MOT) studies. It seems likely that groups of spatially separate entities undergoing common motion are construed as individuals in these studies. As we argue in Section 6, the infant literature bears on this issue.

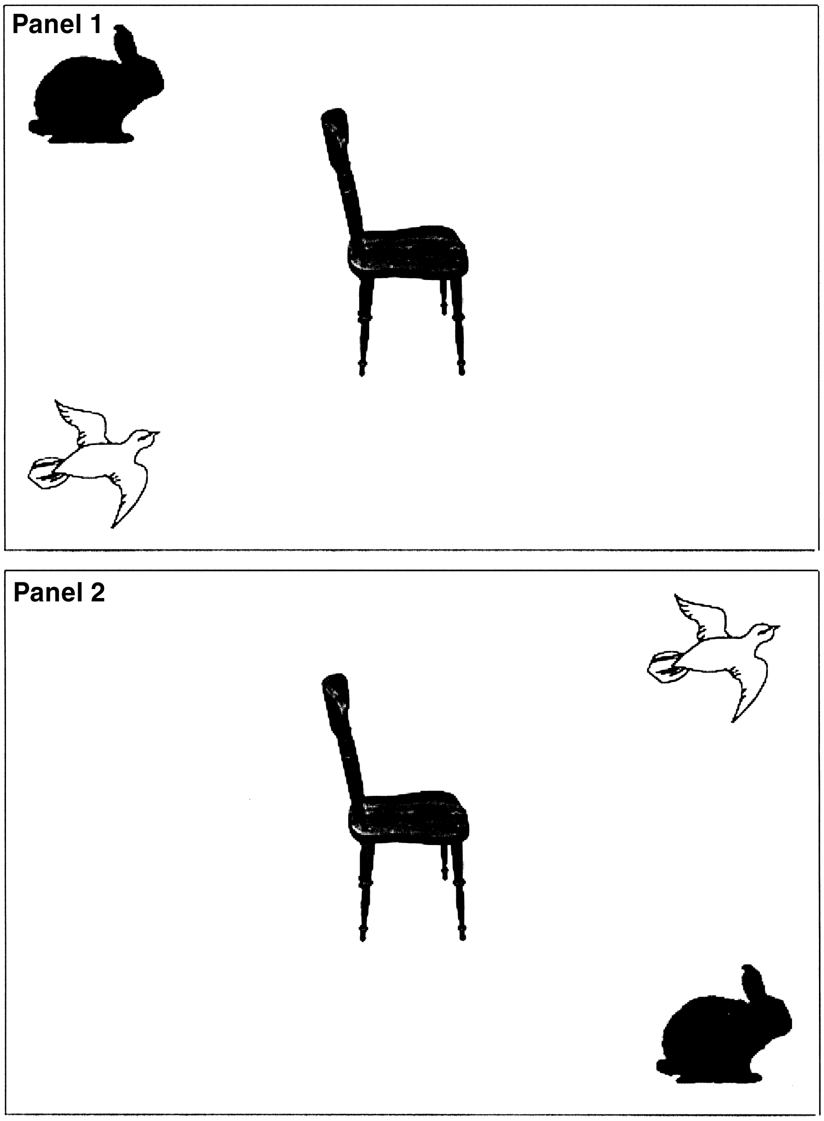

Fig. 1. *Prima facie* evidence for two mechanisms of object individuation.

individual as a large cup a month later. Similarly, color differences do not signal distinct individual chameleons, but they do signal distinct individual frogs.

Fig. 1 illustrates the operation of the two systems in establishing numerical identity. First examine Panel 1. Imagine that you lose perceptual contact with the scene, and return 5 min later to view Panel 2. How would you describe what has happened? You would probably say that the rabbit has moved from above and to the left of the chair to below and to the right of it, while the bird has moved from the bottom left to the top right. In this account, numerical identity (sameness in the sense of *same one*) is being carried by kind membership; it is the rabbit and the bird each of whom you assume has moved through time. The conceptual, kind-based, system of individuation is responsible for establishing the object tokens in this case. Now imagine that a fixation point replaces the chair, and Panels 1 and 2 are projected one after the other onto a screen, while you maintain fixation on the common fixation point. If the timing of the stimuli supports apparent motion, what would your perception be? Rather than seeing a bird and a rabbit each moving diagonally, you see two individuals each changing back and forth between a white bird-shaped object and a black rabbit-shaped object as they move side to side. The visual system that computes the numerical identity of the objects that undergo apparent motion in arrays such as Fig. 1 minimizes the total amount of movement; this system takes into account property or kind information only when spatiotemporal considerations are equated (see Nakayama, He, & Shimojo, 1995, for a review). The mid-level object file system is responsible for establishing the object tokens in this case, and it settles on a different solution than does the kind-based object individuation system.

We shall argue that studies on object individuation in infancy lend support for the suggestion that kind-based object individuation is architecturally distinct from the mid-level object file system. But we must begin by providing some evidence that, contra Piaget, young infants establish representations of individual objects and track them through time. Before we consider the nature of the processes that subserve object representations in early infancy, we must be convinced that there *are* object representations in early infancy.

3. Object individuation and numerical identity in the first year of life

Studies using the violation of expectancy looking time methodology have pushed back the age of the representation of object permanence to 2.5 months (Baillargeon & DeVos, 1991; Hespos & Baillargeon, in press; Spelke, Breinlinger, Macomber, & Jacobson, 1992). In these experiments, infants watch events unfold before them. After being familiarized or habituated to the events, typically they are shown, in alternation, an expected outcome (an outcome that is consistent with adults' understanding of the physical world) and an unexpected outcome (an outcome that is inconsistent with adults' understanding of the physical world, a magic trick). If infants have the same understanding of the events as do adults, they should look longer at the unexpected outcome relative to the expected outcome. Often, but not always, these studies involve events unfolding behind screens, the outcome of the magic trick being revealed upon removal of the screen. These studies require no training; one simply monitors looking times as the infant watches what is happening. Thus, this method taps spontaneous representation of objects and events.

This method yields interpretable findings in newborns (e.g. Slater, Johnson, Brown, & Badenoch, 1996), and is widely used in studies of infants of 2 months

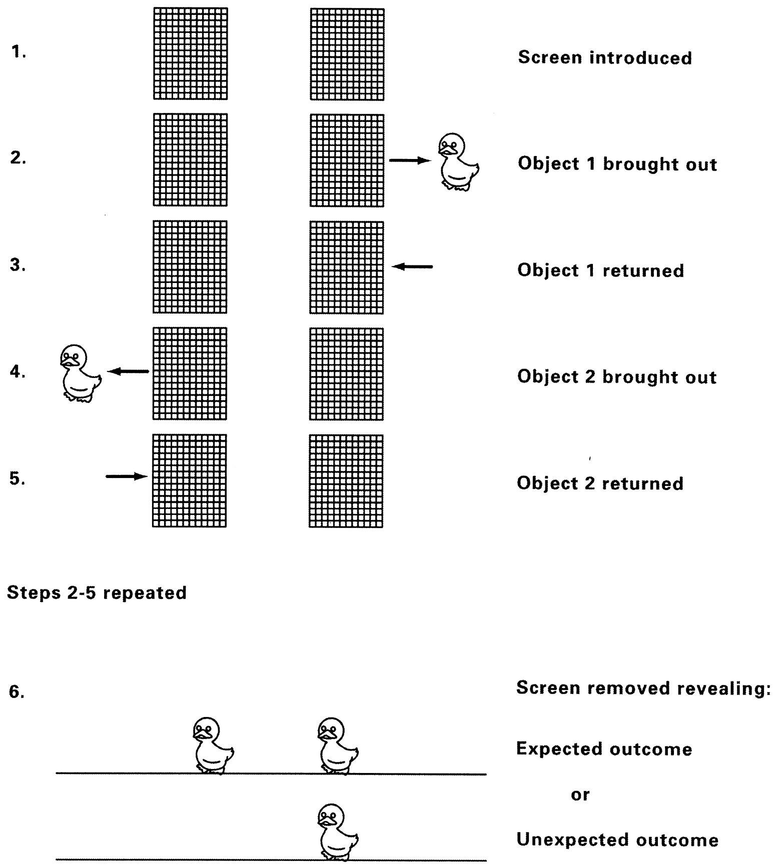

Fig. 2. Schematic representation of experimental paradigm in Spelke, Kestenbaum, Simons, and Wein (1995).

and older. Here we briefly describe two studies using this methodology that illuminate the relation between object permanence and infants' use of spatiotemporal information in the service of object individuation. By spatiotemporal information we mean location or motion information – spatial separation in the frontal plane or in depth, and continuity or discontinuity in an object's trajectory.

Spelke, Kestenbaum, Simons, and Wein (1995) showed that infants do not merely expect objects to continue to exist when out of view, but also that they interpret spatiotemporal discontinuity as evidence for two numerically distinct objects. They showed 4.5-month-old infants two screens with a gap in between, from which objects emerged as in Fig. 2. One object emerged from the left edge of the left screen and then returned behind that screen, and after a suitable delay, a second, physically identical object emerged from the right edge of the right screen and then returned behind it. No object ever appeared in the space between the two screens.

Since an object cannot get from point A to point B without traversing a spatiotemporally continuous path, adults conclude that there must be two numerically distinct objects involved in this event. What about these young infants? After habituation, the screens were removed, revealing only one object (the unexpected outcome) or two objects (the expected outcome). The infants looked reliably longer at the one-object outcome, suggesting they, too, established representations of two distinct objects in this event. A control condition established that infants indeed analyzed the path of motion, and did not expect two objects just because there were two screens. If the object did appear in the space between the two screens, a different pattern of looking was obtained.[2]

Using a different procedure, Wynn (1992) provided further evidence that infants are able to use spatiotemporal discontinuity in object individuation. Five-month-old infants watched a Mickey Mouse doll being placed on a puppet stage. The experimenter then occluded the doll from the infant's view by raising a screen, and placed a second doll behind the screen. The screen was then lowered, revealing either the expected outcome of 2 dolls, or the unexpected outcome of 1 doll or 3 dolls. Infants looked longer at the unexpected outcomes of 1 or 3 objects than at the expected outcome of 2 objects. Wynn interpreted these studies as showing that infants can add $1 + 1$ to yield precisely 2.[3] Whatever these studies tell us about infants' capacity for addition, success depends on the infant's ability to use spatiotemporal discontinuity to infer that the second Mickey Mouse doll was numerically distinct from the first one.

These results suggest that (1) infants represent objects as continuing to exist when they are invisible behind barriers, (2) infants distinguish one object from two numerically distinct but featurally identical objects, distinguishing *one object* from *one object and another object*, and (3) the information infants draw upon for object individuation and for establishing numerical identity is spatiotemporal. If spatiotemporal discontinuity is detected, young infants establish representations of two numerically distinct objects.

Contrary to Piaget's position that processes for establishing representations that trace individual objects through time and occlusion develop slowly over the first 2 years of human life, these studies indicate that they are in place by 4 months of age.

[2] In Spelke, Kestenbaum, Simons, and Wein (1995), 5-month-old infants were agnostic as to how many objects were involved in the continuous event; in a replication with 10-month-olds, infants established a representation of a single object in the continuous motion event, reversing the pattern of preference from the outcomes of the discontinuous motion condition (Xu & Carey, 1996). What is important is that in both experiments the pattern of looking differed between the two conditions (continuous motion vs. discontinuous motion). Thus, when they detected spatiotemporal discontinuity, infants created representations of two numerically distinct, though featurally identical, objects.

[3] Wynn (1992) and her many replicators (Feigenson, Carey, & Spelke, in press; Koechlin, Dehaene, & Mehler, 1998; Simon, Hespos, & Rochat, 1995) also included a subtraction condition: $2 - 1 = 2$ or 1. Infants looked longer at the outcomes of two objects, the unexpected outcome in this condition. Irrespective of what these studies show about infant representation of number (see Simon, 1997; Uller et al., 1999), here we emphasize their implications for infant representations of objects.

Other studies push this age as low as 2 months (e.g. Hespos & Baillargeon, in press), and some have argued that these abilities may be given innately (e.g. Spelke, 1996).

4. Does the mid-level object file system underlie infant object representations?

As argued by Leslie et al. (1998) and Scholl and Leslie (1999), the identification of object representations in young infants with the object files of object-based attention rests on several considerations. First, and most importantly, both systems privilege spatiotemporal information in decisions about individuation and numerical identity. Second, both systems are subject to the same set size limitation of parallel individuation; that is, only three (or four) objects can be indexed and tracked simultaneously. Third, the object representations of both systems survive occlusion, and object tracking is sensitive to the distinction between loss of visual contact that signals cessation of existence and loss of visual contact that does not.

4.1. Primacy of spatiotemporal information

In the mid-level object file system, the questions of individuation and numerical identity concern the bases on which an indexed object retains its index, as opposed to a new object file being established or a new index being assigned. Pylyshyn (2001) and Scholl (2001) both touch on evidence suggesting that spatiotemporal continuity is the primary determinant of numerical identity in this system. Features of an indexed object can change and may be represented as such (see also Kahneman et al., 1992). This is seen clearly in apparent motion studies; the visual system has no problem seeing totally distinct features as states of a single moving object. In order to see apparent motion in cases such as that illustrated in Fig. 1, the visual system must decide which object to pair with which object. To a first approximation, spatiotemporal considerations decide the matter. In such simple displays, the system will minimize the total amount of movement, and will happily override featural information in favor of a motion of two objects each changing color, size and shape as well as kind. However, featural information can play a secondary role: when spatiotemporal information does not unambiguously favor one solution over the other, featural changes are taken into account (see Nakayama et al., 1995, for a review).

The phenomenon of the “tunnel effect” (Burke, 1952) further underscores that new object files are not opened on the basis of featural differences. The tunnel effect is the perception of object unity when objects disappear behind a barrier, reappearing later out the other side. Michotte and Burke (1951) dubbed this phenomenon “amodal completion” because observers do not *see* the object behind the screen (unlike in apparent motion or subjective contours). Rather, observers encode the event as involving a single object despite the discontinuity of perceptual input, and they can even describe its hidden trajectory. Spatiotemporal considerations determine amodal completion (the speed of the object, the time behind the occluder, the relative sizes of the objects to that of the occluder). What do *not* matter are the features of the objects; a green circle entering behind the screen may emerge as a red

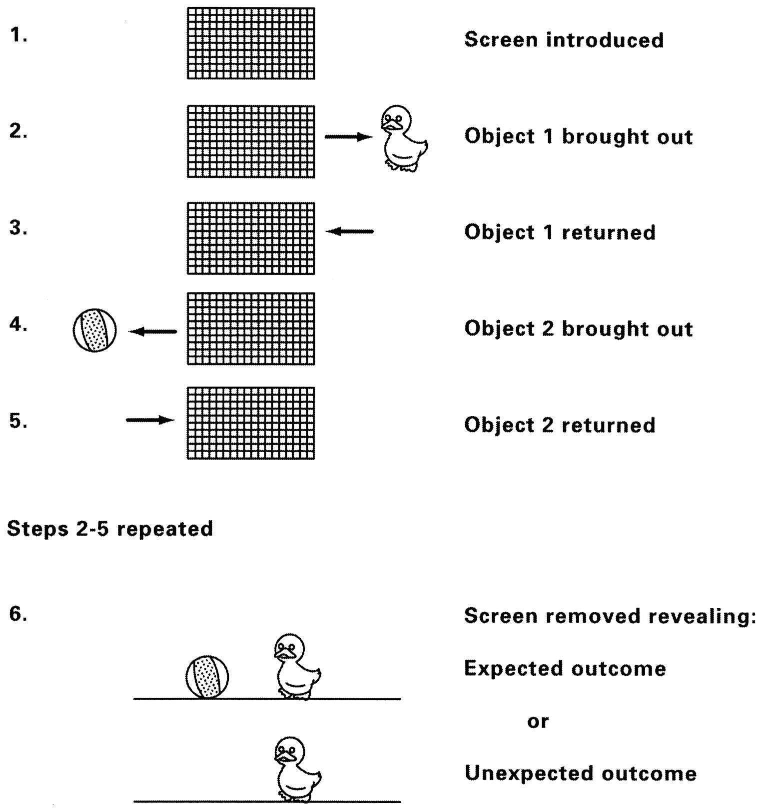

Fig. 3. Schematic representation of experimental paradigm from Xu and Carey (1996).

square and yet be seen as the same object just as strongly as if it emerges a green circle, so long as the spatiotemporal parameters supporting amodal completion are met (Burke, 1952).

Consistent with the claim that featural changes do not signal the opening of new object files, object tracking in the MOT studies is not disrupted by indexed objects changing color, size, shape or kind during tracking (Pylyshyn, 2001). Finally, a recent study by Scholl, Pylyshyn, and Franconeri (2001) underscores the primacy of spatiotemporal information in the establishing and tracking of object files. In the MOT paradigm, if tracking is stopped and one of the objects disappears, the subjects can indicate its location and direction of motion. But if objects are changing properties during tracking, subjects are not aware of the momentary color or shape of a tracked object.

In sum, the computations that maintain indexes to attended objects rely heavily on spatiotemporal information; objects are tracked on the basis of spatiotemporal continuity. Once an object file is opened, features may be bound to it, and updated as the object moves through space. (The study just described shows that features are not automatically bound to open object files, perhaps because of the high attentional demands of tracking four independently moving objects at once.) These generalizations hold for the young infant's object representations as well, the point to which we now turn.

The Spelke, Kestenbaum, Simons, and Wein (1995) and Wynn (1992) studies described above suggest that infants as young as 4 months of age draw on spatiotemporal information in object individuation and tracking, but they do not show that spatiotemporal information is privileged, for they did not explore whether infants could also use property or kind differences as a basis for object individuation. Recent studies suggest that young infants do not use property or kind differences as a basis for opening new object files (Xu & Carey, 1996), especially when spatiotemporal evidence is strong (e.g. continuous trajectory specifying one object, a single location specifying one object). Imagine the following scenario. One screen is put on a puppet stage. A duck emerges from behind the screen and returns behind it, and then a ball emerges from behind the same screen and then returns (Fig. 3). How many objects are behind the screen? For adults, the answer is clear: At least two, a duck and a ball. But since there is only a single screen occluding the objects, there is no clear spatiotemporal evidence that there are two objects. We must rely on our knowledge about object properties or object kinds to succeed at this task. In our studies, infants were shown the above event. The contrast was either at the superordinate (as well as basic) level (e.g. a duck and a ball, an elephant and a truck; or an animal and a vehicle) or just at the basic level (e.g. a cup and a ball); some objects were toy models (e.g. truck, duck) where others were from highly familiar everyday kinds (e.g. cup, bottle, book, ball). On the test trials, the screen was removed to reveal either the expected outcome of two objects or the unexpected outcome of only one of them. If infants have the same expectations as adults, they should look longer at the unexpected outcome. The results, however, were surprising: 10-month-old infants failed to draw the inference that there should be two objects behind the screen, whereas 12-month-old infants succeeded in doing so.

Control conditions established that the method was sensitive. Ten-month-old infants succeeded at the task if they were given spatiotemporal evidence that there were two numerically distinct objects, e.g. if they were shown the two objects *simultaneously* for 2 or 3 s at the beginning of the experiment. Furthermore, Xu and Carey (1996) showed that infants are sensitive to object properties under the circumstances of their experimental paradigm: it takes infants longer to habituate to a duck and a car alternately appearing from behind the screen than to a car repeatedly appearing from behind the screen. In this task, infants failed to draw on object kind information for object individuation (e.g. animal, vehicle, duck, truck, ball, cup, etc.); they also failed to draw on property contrasts (e.g. the contrast between being yellow, curvilinear, and rubber vs. being red, rectilinear, and metal). The property differences which infants under 10 months of age are sensitive to may be irrelevant to object individuation. Other laboratories have replicated these findings (Wilcox & Baillargeon, 1998a, Experiments 1 and 2; see Xu & Carey, 2000, and Section 5.2 below, for a discussion of some apparently conflicting data from Wilcox, 1999; Wilcox & Baillargeon, 1998a,b).

Van de Walle, Carey, and Prevor (in press) sought convergent evidence for the claim that infants below 12 months of age do not use kind membership as a basis for opening new object files. In these studies, a manual search measure was used instead of the violation of expectancy looking time procedure. Ten- and 12-month-old infants were trained to reach through a spandex slit into a box into which they could not see in order to retrieve objects. Three types of trials were contrasted: one-object trials, two-object trials in which individuation must be based on property/kind contrasts, and two-object trials in which spatiotemporal evidence specified numerically distinct objects. On a one-object trial, the experimenter pulled out the same object (e.g. the toy telephone) twice, replacing it into the box each time. On two-object trials in which individuation is based on property/kind information, infants watched the experimenter pull out an object (e.g. a toy telephone), return it to the box, then pull out a second object (e.g. a toy duck), and return it to the box. On two-object trials in which spatiotemporal evidence supported individuation, the experimenter pulled out the first object (e.g. the telephone), left it on top of the box, pulled out the second object (e.g. the duck) so that they were simultaneously visible, and then returned both to the box.

The boxes were then pushed into the child's reach, and patterns of search revealed how many objects the child had represented as being in the box. Both 10- and 12-month-olds differentiated the one- and two-object trials when given spatiotemporal evidence for two objects. That is, they searched for a second object after having retrieved the first one on two-object trials but not on one-object trials, and having retrieved the second object on two-object trials, they did not search further. Twelve-month-olds also succeeded when given property/kind information alone. In contrast, the 10-month-olds failed in this condition; their pattern of search on the two-object trials was the same as on the one-object trials. Ten-month-olds failed to use kind differences such as telephone, duck or car, book or property differences such as black, yellow, telephone-shaped, duck-shaped, rubber, or plastic to establish representations of two numerically distinct objects in the box. These results are consistent with those of the looking time studies of Xu and Carey (1996).

We draw two conclusions from these studies. First, they support the identification of the young infants' object representations with those of the mid-level object file system, for they show that infants under 10 months of age rely almost exclusively on spatiotemporal information in decisions about numerical identity of objects seen at different times. Second, they are consistent with the possibility that a second system of object individuation, a kind-based system, emerges at around 12 months of age (see Section 6.1 for further discussion).

4.2. Set size limitations

Pylyshyn's MOT paradigm provides direct evidence regarding the number of objects that may be simultaneously indexed and tracked through time. Although various task variables affect the set size at which performance is virtually errorless, a good approximation is that about four objects are the limit (see Pylyshyn, 2001; Trick & Pylyshyn, 1994, for a discussion of the relations between the limits on parallel individuation and indexing of objects and the limits on subitization, the rapid apprehension of precise numerosity of small sets of objects, in the absence of counting).

Results from several experimental paradigms suggest that young infants' limit on parallel individuation of objects is in the same range. In the interest of space, we mention just two lines of relevant work. The studies by Spelke, Kestenbaum, Simons, and Wein (1995) and Wynn (1992), described above, show that infants represent events in terms of precisely one object or precisely two objects. Success with sets of three objects, however, is mixed: Wynn (1992) showed that 4-month-old infants expected $1 + 1$ to be precisely 2 (they looked longer at impossible outcomes of 3 than at possible outcomes of 2, as well as at impossible outcomes of 1). Wynn also found that young infants succeeded at a $3 - 1 = 2$ compared to a $2 + 1 = 2$ comparison. Baillargeon, Miller, and Constantino (1993) found that 10-month-olds succeeded in a $2 + 1 = 3$ or 2 comparison, but they failed at a $1 + 1 + 1 = 3$ or 2 comparison. Finally, Uller and Leslie (1999) found that 10-month-olds succeeded in a $2 - 0 = 1$ vs. $2 - 1 = 1$ comparison, but failed in a $3 - 0 = 2$ vs. $3 - 1 = 2$ comparison. Thus, there appears to be robust successes with sets of 1 and 2, and some fragile successes with sets of 3.

Similarly, in simple habituation paradigms, in which, over time, infants look less at successive presentations of arrays of a single set size (e.g. 3) and recover interest when shown an array of a different set size (e.g. 2), performance often falls apart at 3 vs. 4 (Starkey & Cooper, 1980). That parallel individuation of small sets of objects underlies success in these studies, rather than a symbolic representation of number such as that computed by analog magnitude systems (Dehaene, 1997), shows that success is not predicted by Weber fraction considerations; infants succeed at 2 vs. 3 but fail at 4 vs. 6 (e.g. Starkey & Cooper, 1980).[4] Thus, that the limits on set sizes of object tokens that may be simultaneously attended and tracked are in the same range supports the identification of the system that supports object individuation in infancy with that underlying object-based attention in adults.

4.3. Occlusion vs. existence cessation

Another parallel between the two systems is that indexed objects, just like the objects represented by infants, survive occlusion, as revealed in studies of the tunnel effect (Burke, 1952). Further, Scholl and Pylyshyn (1999) showed that object tracking in the MOT paradigm was not disrupted by the objects going behind real or virtual occluders. Almost all of the infant studies cited above involve occlusion.

In Scholl and Pylyshyn (1999) it mattered that the objects disappeared behind an occluder by regular deletion along its contour, reemerging from the other side by regular accretion along its opposite contour. If the objects disappeared at the same rate by shrinking to nothing, reappearing farther along the trajectory at the same rate

[4] Although experiments with small sets of objects reveal the set size signature of object file representations (Feigenson et al., 2001), under some circumstances infants also create numerical representations of large sets that show the Weber fraction signature of analog magnitude representations (Xu, 2000; Xu & Spelke, 2000).

by expanding from a point, tracking was totally disrupted. Thus, the system distinguished the object's going behind an occluder from its going out of existence, to be later replaced by another object coming into existence. Bower (1974) provided evidence that young infants draw the same distinction. Bower compared infants' visual search for objects that disappeared by shrinking down to nothing with their visual search for objects that disappeared by progressive deletion along a boundary. Infants searched for the missing object in the latter case but not the former. This early experiment bears replication, perhaps with a manual search paradigm along the lines of Van de Walle et al. (in press).

4.4. Conclusions

Section 4 has outlined the considerations in favor of identifying infants' object representations and object files, as well as identifying the computations that underlie young infants' tracking of moving with the adult mid-level system of object indexing. For the rest of this review, we will adopt this identification as a working hypothesis, and consider its implications for each of the two research communities.

What is to be gained from the discovery that students of adult mid-level object-based attention and students of infant object representations are exploring the same natural kind? Some have argued that this discovery *explains* some of the properties of infant object representations, such as the primacy of spatiotemporal information in individuation or the set size limitations. Of course, this is not so; at best, the identification reduces two sets of mysteries to one. Still, both communities stand to benefit from this discovery. Understanding hard won in one community may be applied to the other, and phenomena explored in one literature become a source of hypothesis for the other.

5. Lessons to be learned regarding infants' object representations

5.1. Object representations in infancy: perceptual or conceptual?

As Scholl and Leslie (1999) discussed at length, that infant object representations are object files has important implications for the controversies in the infant literature concerning whether infants' object representations are *conceptual* or *perceptual*. In the attention literature, object file representations are considered *mid-level* between low-level sensory representations and fully conceptual representations. Object file representations do not depend upon categorizing individuals into antecedently represented object kinds. To a large extent, the mechanisms that index and track objects through time work the same way whether the objects are instances of familiar kinds or not (see Nakayama et al., 1995, for a review), and are thus *mid-level* in not requiring placement into conceptual categories.

Scholl and Leslie (1999) had a different sense of *mid-level* in mind. They were concerned with the status of the spatiotemporal and featural information that enters into the processes of object indexing and object file creation. It is consistent with their position that the object files themselves are symbolic representations (see

Sections 6.4–6.7 below). Nonetheless, spatiotemporal and featural information that is drawn upon in the creation and maintenance of object files is most likely represented in an encapsulated perceptual system (see Pylyshyn, 2001). If so, it is misleading to say that the infant "believes" that objects trace spatiotemporally continuous paths, or "knows" that objects are permanent, for the infant represents no such propositions in any accessible form. We are in agreement with Scholl and Leslie, and with Pylyshyn, on this point.

5.2. Object featural information and the tunnel effect

The identification of the two literatures is a source of insight into the different status of spatiotemporal information and object feature information in the young infant's object representations. However, it is controversial that spatiotemporal information takes precedence over featural information in infants' individuation of objects (Needham & Baillargeon, 1997; Wilcox, 1999; Wilcox & Baillargeon, 1998a,b). This controversy potentially undermines the identification of the two literatures. However, when we look more closely at these apparent conflicts, uniting the two literatures helps us resolve them, and thereby strengthens the integration.

A central piece of evidence for the identification of the two literatures is the failure of infants to draw on featural differences in establishing representations of two objects in the studies of Van de Walle et al. (in press) and Xu and Carey (1996) described above. Recent studies by Wilcox and her colleagues (Wilcox & Baillargeon, 1998a,b) have challenged our interpretation of these results. In Wilcox and Baillargeon's narrow/wide-screen studies, infants watched a blue ball and a red box emerge, one at a time, from opposite sides of a screen. In each cycle, both objects were out of view, behind the screen, for a short period of time. Two conditions were contrasted. In the wide-screen condition, the occluding screen was 30 cm wide, wide enough for both objects to simultaneously fit behind, since the sum of the widths of the ball and the box was 22 cm. In the narrow-screen condition, however, the screen was too narrow (21 cm or even narrower) for both objects to fit behind. Infants as young as 4.5 months of age looked longer at the narrow-screen event than at the wide-screen event. Wilcox and Baillargeon interpreted the infants' behavior as follows: in the narrow-screen event, the infants must have used the featural (or kind) differences between the box and the ball to infer that two distinct objects were involved in this event and must have realized that the two objects could not fit behind the screen simultaneously.

These are extremely creative and interesting studies. However, there is another possible interpretation of the results. The narrow/wide-screen events are very similar to those in studies of the tunnel effect described above (e.g. Burke, 1952). In amodal completion, the visual system takes into account various spatiotemporal parameters and yields a representation of a single object persisting through occlusion. Perhaps the conditions of the narrow-screen event are those that support amodal completion, such that the infant represents it as a single object persisting through occlusion, and finds the change of properties anomalous. Although babies, by hypothesis deploying the mid-level object tracking system, can update representations of single objects when properties change, they nonetheless expect an object's properties to stay constant.[5] On this alternative account, infants are not using the property differences as a basis for opening a second object file in the narrow-screen events; rather the property change of a single object is anomalous, and thus attention grabbing. On this account, the wide-screen events do not yield amodal completion so there is no single object-token whose properties changed during occlusion.

To explore the amodal completion hypothesis, Carey and Bassin (1998) assessed adults' spontaneous perception of the events upon seeing them (without any verbal prompting, a situation identical to what the infant experienced). Virtually all of the participants shown a very narrow-screen (15 cm) event spontaneously noted that something was anomalous, as did 40% of those shown a 21 cm narrow-screen event. Most importantly, *all but one* of the participants, when they noticed the anomaly, whether in the 15 cm or the 21 cm version, described it as follows: "It went in a ball and it came out a box." That is, they described the event as a single object magically changing properties (as described in the tunnel effect literature), rather than two objects that could not fit behind the screen.

Notice that the tunnel effect alternative interpretation assumes that infants, like adults, used the relative size of the objects and the occluder to establish a representation of a single object persisting behind the screen, and that infants, like adults, expect that properties of objects remain constant while occluded, and thus find the property changes interesting or anomalous. On this interpretation, the developmental changes reported in Wilcox (1999) concern *which* property changes of a single object infants find anomalous or interesting (first size and shape, then surface pattern, then color). On this interpretation, the narrow-screen findings do *not* reflect the child's ability to use featural information as a basis for decisions of numerical identity of object files.

It is, of course, an open question whether our interpretation of the narrow-screen/wide-screen studies is correct. We offer it here as an example of how the identification of the two literatures might guide the interpretation of apparently conflicting results. Furthermore, our hypothesis suggests experiments on the tunnel effect in adults. To our knowledge, there has been no systematic study of the effects of the relative size of the objects and the occluders in producing an illusion of a single object behind the barrier. The screens in the adult studies of the tunnel effect are much wider, relative to the objects, than those in these infant studies. The Carey and Bassin (1998) findings should be systematically followed up; the conditions of the narrow-screen events should produce very strong amodal completion, irrespective of object speed.

[5] There is ample evidence that infants expect properties bound to a represented object to remain constant during occlusion. For example, in Baillargeon's rotating screen studies, infants predict when the screen's motion should be arrested from the height of the occluded object (Baillargeon, 1991), and in Aguilar and Baillargeon's studies of when objects should be visible after going behind screens with a window, infants again take into account the height of the occluded object (e.g. Aguilar & Baillargeon, 1999).

5.3. A second challenge to the Xu and Carey (1996) findings

Experiments 7 and 8 of Wilcox and Baillargeon (1998a) show that young infants use featural information for object individuation, and are not subject to the tunnel effect interpretation. In their study, 9.5-month-old infants were shown a box moving from one side of the stage and disappearing behind a screen, followed by a ball emerging from the other side of the screen. The screen was then lowered and the infant saw *only* the ball on the stage. Infants looked longer at this outcome relative to a condition where the same ball disappeared behind the screen and reappeared from the other side. However, this positive result goes away completely if the first object, the box, appeared from behind the screen, moved to the side of the stage, then reversed its trajectory and disappeared behind the same screen, the ball then emerging from behind the same screen. The test outcome was identical to the experiment described above.

Wilcox and Baillargeon (1998a) argued that the infants' success in the first condition is due to their using the differences between the box and the ball to create representations of two objects, their attention being drawn by the anomalous ball only outcome in the ball–box condition. We agree with their argument. One possible interpretation for the success in the single trajectory condition, in the face of failure in the double trajectory condition (as well as in Van de Walle et al., in press; Xu & Carey, 1996), is that the single trajectory condition provided very little spatiotemporal information that there was a single object. Analogous to the case of apparent motion, when spatiotemporal evidence does not favor one solution over another, infants can use featural differences for object individuation. However, slightly stronger spatiotemporal evidence for the presence of a single object (as in the second experiment with a reversal of trajectory and both objects appearing from behind the same location, namely the screen) overrides any sensitivity to features and the object file system computes a representation of a single one object. In the Xu and Carey (1996) studies, spatiotemporal evidence for one object was even greater; the objects emerged from behind the same screen several times, and reversed trajectory several times.

5.4. Object segregation vs. object files

Xu, Carey, and Welch (1999) explored when infants could use feature or kind information to individuate objects in static arrays, and found age shifts that converged with those found in the individuation within object tracking experiments cited above. Consider Fig. 4. How many objects are there in this array? Adults respond that there are two objects, a duck and a car, and if the duck is lifted from above, adults predict that the duck will come alone and are surprised if the duck/car moves as a single object. Xu et al. (1999) habituated 10- and 12-month-old infants to the stationary duck/car stimulus (and to an analogous cup/shoe stimulus), after which the top object was grasped and lifted, and two outcomes shown in alternation. In one outcome, just the top object came up (move-apart outcome) and in the other, both objects came up (move-together outcome). At 10 months, the infants did not look longer at the unexpected, move-together, outcome. They failed to use the

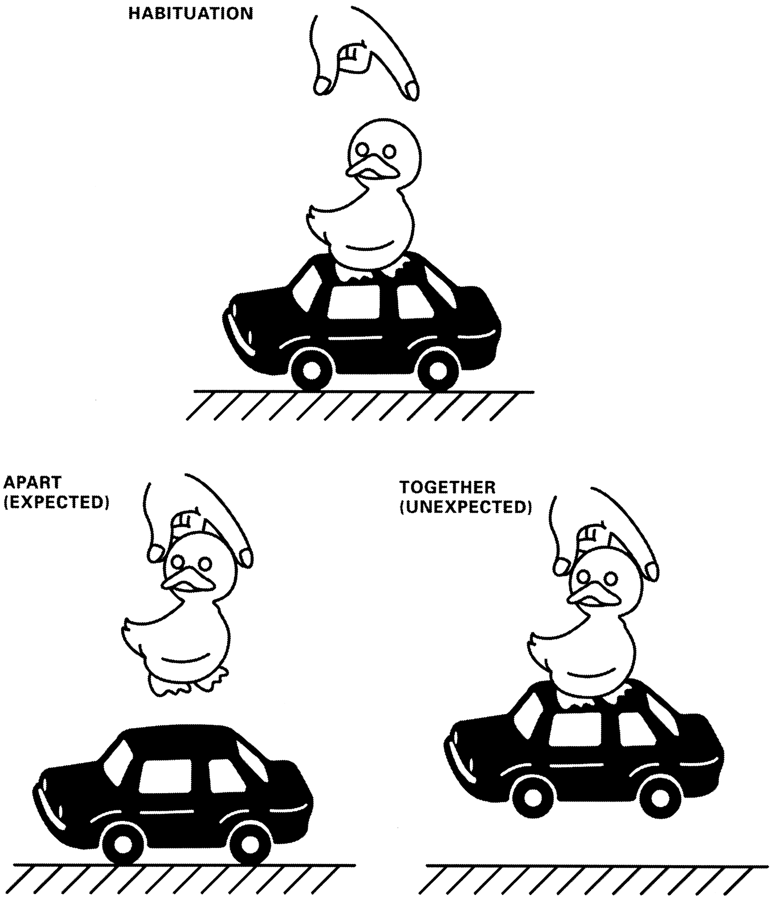

Fig. 4. Schematic representation of experimental paradigm from Xu et al. (1999).

contrast between the duck and the car, or the cup and the shoe, to infer that there were two individual objects in the array. That is, 10-month-olds failed to draw on kind contrasts (duck/car, cup/shoe) or property contrasts (yellow-rubber-duck shaped/red-metal-car shaped) to resolve the ambiguous object into two. At 12 months, however, the infants succeed at the task, looking longer at the unexpected outcome in which the cup/shoe or duck/car moved as a single object. Furthermore, as in Van de Walle et al. (in press) and Xu and Carey (1996), when 10-month-olds were given spatiotemporal evidence that there were two objects (e.g. if the objects were briefly moved, laterally, relative to each at the beginning of each habituation trial), they now succeeded, looking longer at the unexpected move-together outcome. These results converge with the data of Van de Walle et al. (in press) and Xu and Carey (1996). However, these results are in apparent conflict with other experiments by Needham and her colleagues.

Needham and her colleagues (Needham, 1998; Needham & Baillargeon, 1997,

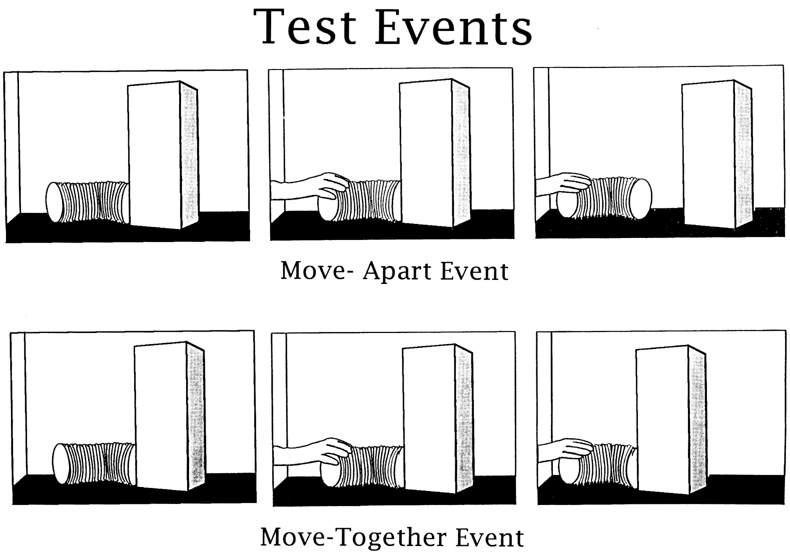

Fig. 5. Schematic representation of ambiguous stimulus from Needham box/hose object segregation studies.

1998) demonstrated that even infants as young as 5 months of age succeed in using featural information to segment objects that share boundaries. Consider Fig. 5. The rectangular box is blue and made of wood, and the cylinder is bright yellow and made of plastic. Young infants use the contrast between blue, rectangular, wood and yellow, cylindrical, plastic, or some subset of these contrasts, to resolve the ambiguity derived from a shared boundary and to parse this figure into two distinct objects. This is demonstrated in experiments in which infants view this ambiguous display for a few seconds, after which one of the objects is grasped and pulled. Infants look longer if the box/hose object moves as a single whole than if it comes apart.

Thus, featural information plays a role in object segregation problems in early infancy, but this does not undermine the arguments of Section 4.1 above, for the processes of object indexing, creating object files, and tracking objects through time engage two quite distinct individuation problems. First, there is the segregation of objects that share boundaries, which, like figure/ground segregation, concerns the problem of assigning edges and surfaces to individuals. In these problems, such as those posed by the displays in Figs. 4 and 5, ambiguity arises from shared boundaries. Second, there is the *different* individuation problem than that which arises in object tracking experiments. Object tracking experiments concern, among other things, whether already perceptually segregated individuals are numerically distinct. Ambiguity in the latter case arises because perceptual contact specifying spatiotemporal continuity has been lost (as in occlusion, or due to attentional shifts). As we have indicated at length, it is in the latter problem that featural information plays a decidedly secondary role. But in the former cases (figure/ground and object segregation in static arrays), featural information *must* play a primary role, for edges are specified by color and brightness contrasts. Other featural cues, such as gestalt cues (good form, symmetry, feature similarity) also enter into the earliest stages of figure/ground segregation, as does spatiotemporal information such as spatial segregation in depth. It is very likely that all of these cues would influence infants' object segregation as well, an empirical matter worth exploring.

In sum, the adult literature distinguishes the processes through which edges are assigned to figures, or surfaces to objects, on the one hand, from those through which already segmented figures or objects are tracked through time, their features updated as they change. In the former processes, featural information plays a pivotal role, in part with and in interaction with spatiotemporal information, while in the latter spatiotemporal information is sharply privileged. Thus, that young infants (at least as young as 4 months of age) make robust use of featural information for object segregation does not undermine the claim that they almost always fail to do so in the service of tracing numerical identity of already segmented objects.

Why then did infants in Xu et al. (1999) fail to use the distinctions between the duck and the car, or the cup and shoe to segment the arrays as in Fig. 4 into two objects? We recruit two further distinctions in our speculative answer to this question.

5.5. On distinguishing featural/property information from kind information

The merging of the two literatures supports another distinction that might help resolve some of the apparent empirical conflicts in the infant literature. Recall that very young infants succeed at segmenting the ambiguous box/hose arrays (Fig. 5) on the basis of featural differences between the two objects, but it was not until 12 months that infants segmented the ambiguous duck/car display (Fig. 4) into two objects. Needham and colleagues have found that success at any given object segmentation task is sensitive to object complexity. For instance, making the hose curved and rotating the array so that the boundary between the box and hose isn't fully visible pushes the age of success a few months older (Needham, Baillargeon, & Kauffman, 1997).

The duck/car and cup/shoe stimuli of Xu et al. (1999) were more complex than any that have been used in the Needham et al. studies. They are multicolored and multi-parted, with each object having a complex, irregular shape. Their properties alone do not support an unambiguous parse; property contrasts support segmenting the head from the body, the beak from the rest of the head, the body from the feet, the windowed part of the car from the rest of the car, the wheels from the rest of the car, as well as the duck from the car.[6]

[6] Needham and Baillargeon (2000) and Xu and Carey (2000) discuss many other respects in which the Xu et al. (1999) paradigm poses a more difficult problem for infants than does the paradigm used by Needham and her collaborators. For example, babies in our studies are habituated to the stationary display, perhaps supporting an interpretation of the array as a single object. Also, Needham and Baillargeon (2000) review unpublished work that shows that infants succeed in segmenting side by side objects at a younger age than they do objects one on top of the other. We suggest that each of these factors makes it more likely that infants need to draw on kind representations to solve the problem, and that it is kind representations that are becoming available between 10 and 12 months of age.

It is important to distinguish the encapsulated processes that draw on property information in object segregation from processes that draw on conceptually mediated kind representations, as in recognizing the top part of Fig. 5 as a duck (see Pylyshyn, 2001, for an extended discussion of this distinction). Xu and Carey (2000) suggest that various features of our task, including the fact that the property differences do not support an unambiguous parse, make a property-based parse less likely, and require that the child draw on kind representations to succeed. Thus, the 10–12 month shift in these studies may reflect the emergence of kind representations, or the ability to draw upon them in object individuation, just as do the 10–12 month shifts in Van de Walle et al. (in press) and Xu and Carey (1996).

5.6. On distinguishing between kind representations and experience-based shape representations

There is one more apparent conflict in findings between the box/hose experiments of Needham and the duck/shoe experiments of Xu et al. Needham found that at an age at which infants do not succeed at parsing an ambiguous stationary display, a few seconds exposure to one of the objects (e.g. the box alone or the hose alone) before presentation of the ambiguous display leads infants as young as 5 months of age to succeed (see Needham et al., 1997, for a review). That is, early in infancy, experience-based representations may be recruited in the service of object segregation. To check if such prior exposure would help in the duck/car case, Xu et al. (1999) included a condition in which 10-month-olds were given 30 s exposure to the duck alone and 30 s exposure to the car alone, before being habituated to the stationary duck/car display (Fig. 4). Ten-month-old infants still failed. Why would experience help in the box/hose case but not the duck/car case?

The work of Peterson (1994) suggests another distinction we must make in thinking about the representations that play a role in object individuation: representations of kinds and representation of experientially derived shapes. Her work has shown that these two types of representations play distinct roles in the process of figure/ground segregation, suggesting that they might also play distinct roles in the process of object segregation.[7]

In a series of studies, Peterson and her colleagues have studied figure/ground displays in which one of the surfaces is bounded by a meaningful shape (e.g. a face profile or a sea horse) and in which its complement is not. She often manipulates other cues to figure/ground segregation as well (e.g. symmetry, binocular depth cues). What she finds is that meaningfulness of shape (which can only have been derived from experience) enters in parallel with and in interaction with encapsulated perceptual processes at the very earliest stages of figure determination. That is, the meaningful shape is more often seen as figure than its complement, and this

[7] We are not claiming here that the problem of object segmentation and the problem of figure/ground segmentation are one and the same problem, just that they are analogous and should be differentiated from the problem of object identity in tracking experiments.

factor sometimes overrides other cues to figure such as symmetry or depth cues (e.g. Peterson & Gibson, 1993; see Peterson, 1994, for a review).

This state of affairs is perhaps paradoxical. Logically, it would seem that object recognition (place an individual token with respect to an antecedently represented kind) would require prior figure/ground segregation, for one needs the individual to match against stored representations. Peterson resolves the paradox by pointing out that familiarity of shape may enter into the process without requiring that actual recognition (accessing a familiar kind) has taken place. In support of this observation, Peterson, de Gelder, Rapcsak, Gerhardstein, and Bachoud-Levis (in press) presented neuropsychological evidence that the experientially derived shape representations that enter into figure/ground segregation are *not* the kind representations that mediate object recognition. They presented a double dissociation between a visual agnosic patient with bilateral temporal-occipital lobe lesions and a patient with bilateral occipital lesions who was impaired on a variety of sensory and perceptual capacities. Agnosic patients cannot recognize familiar objects; they cannot name them, say what they are for, describe them, or show any other evidence of having placed them with respect to a familiar kind. The agnosic patient nonetheless showed the effects of experientially derived shape on figure determination to an equal extent as normal participants in these studies. That is, she was more likely to see a sea horse as figure than an upside-down sea horse (inversion controls for all other cues to figure/ground segregation), even though she could not recognize the sea horse. The occipital patient showed no effect of experientially derived shape representations in figure/ground decisions, but when he saw the meaningful shape as figure, he could recognize it as well as did normal participants in this experiment.

The Peterson work is relevant to the present discussion because it shows that representations of shape may enter into individuation processes in at least two different ways, only one of which involves recognition with respect to antecedently represented kinds. Although the Peterson work concerns figure/ground segregation, the same may be true for object segregation. As Peterson shows, the representations of shape that enter into the encapsulated early processes are fragmentary and simpler than those that support full-blown object recognition. It is possible that the experientially-based shape representations of the geometrically simple box or hose are influencing these early perceptual processes, and that the child cannot form such representations of the more complex duck or car with so little contact with these stimuli. Continuing along this line of speculation, it may be that only when infants have formed kind representations of ducks and cars does recognition of the objects as members of those categories play a role in the object segregation task posed by the stimulus array of Fig. 4, as well as the numerical identity tasks of Van de Walle et al. (in press) and Xu and Carey (1996).

Thus, in all these cases the adult vision literature on object representations contributes to a *possible* resolution of several apparent conflicts in the infant literature. We suggest that the resolution will depend upon distinguishing between mid-level object file representations, property representations, experience-based shape representations and kind representations, and the respective roles these play in distinct individuation problems (figure/ground segregation, object segregation, object files, and kind-based object individuation).

6. Lessons from the infant literature concerning adult object representations

Section 5 considered lessons gained from the adult literature on mid-level object tracking for our understanding of young infants' object representations. Here we ask how the infant literature can return the favor. What lessons about object individuation in adults might be gleaned from the infant literature?

6.1. Distinguishing object file individuation from kind-based individuation

Until now, we have merely asserted that the adult literature distinguishes kind-based individuation from mid-level object file-based individuation. Actually, the literature is not unequivocal on this matter. On some treatments there is no such thing as kind-based individuation. For example, in the standard treatment of logical form in the literature on formal semantics, "The dog is black" is formalized as "((x) (dog(x) & black(x))". That is, being a dog is a property of an existentially quantified individual picked out in some other way, just as being black is. Similarly, Kahneman et al. (1992) suggest that object files represent individual tokens of objects and that "is a truck" or "is a dog" are features of objects rather than themselves sortals that directly provide criteria of individuation and numerical identity. In support of this way of looking at things, they offer the observation that we can felicitously say, "Its a bird, its a plane, its Superman", referring all along to the same individual.

For reasons beyond the scope of this paper, we do not consider this position tenable (see Carey & Xu, 1999; Macnamara, 1986; Xu, 1997, for discussions of the relevant philosophical literature as it relates to the psychological questions). In adult conceptual life, criteria for individuation and numerical identity are sortal-specific. As mentioned earlier, kind-relevant considerations often override spatio-temporal continuity in judgments of numerical identity. A person, just dead, is not identical to her corpse, in spite of the spatiotemporal continuity of her body. Some philosophers (e.g. Hirsch, 1982) would push this point even further, arguing that not only is spatiotemporal continuity not sufficient for our judgment of identity, but that it is not even necessary. A paradigm example is the following. Suppose you have a watch whose interior needs to be cleaned. You dismantle the watch, scattering the various parts on the desk, then you reassemble the watch after cleaning. During this process, spatiotemporal continuity was lost when the parts were scattered on the desk yet our intuition is clear that when the watch has been reassembled, it is the same watch as the one you started with.

In addition, kinds provide criteria of individuation and numerical identity directly, whereas properties do not. One can count the dogs or the shoes or the fingers in this room, but not the red in this room. Thus, at least as articulated in adult language, kind representations are sharply differentiated from property representations. They are not merely features to be bound to individuals picked out by FINSTs or to individual object files.

The infant literature could bear on this controversy. If it turns out that infant cognitive architecture distinguishes between kinds and properties, and between kinds and object files, the position that these must be distinguished in adult cognitive architecture would receive support. We touched on this suggestion in our attempts to resolve the apparent discrepancies between the box/hose studies (Fig. 5) and the duck/car studies (Fig. 4), but we have not yet really marshaled the evidence for this position. Twelve-month-olds robustly succeed in experiments where individuation is signaled by kind distinctions (Van de Walle et al., in press; Xu & Carey, 1996; Xu et al., 1999). However, it does not follow that *is a duck* has a different status in the 12-month-old's conceptual system than does *is yellow*. The fact that 12-month-olds in these studies formed representations of two objects on the basis of the distinction, for example, between a telephone and a book does not mean that they were using kind representations to do so. After all, adults would assume that a black plastic object was numerically distinct from a red cardboard and paper object, even in the absence of having identified these objects as a telephone and a book.

We have recently completed a series of experiments with 12-month-olds (Xu, Carey, Quint, & Bassin, 2001) to establish whether 12-month-olds' success in our studies was based on property contrasts or kind contrasts. Using the paradigm of Xu and Carey (1996), infants were shown an event in which an object (e.g. a red ball) emerged from behind a screen and returned, followed by an object (e.g. a green ball) emerging from behind the screen from the other side and returning. On the test trials, infants were shown two objects (e.g. a red ball *and* a green ball) or just a single object (e.g. a red ball *or* a green ball) when the screen was removed. We found that even though 12-month-old infants were sensitive to the perceptual differences between the objects, these property changes (i.e. color change alone, size change alone, or the combination of the two) did not lead to successful object individuation. That is, upon seeing a red ball alternating with a green ball (or a big ball and a small ball, or a big red ball and a small green ball), the infant did not conclude that there were two distinct objects behind the screen. In the last experiment of this series, infants were shown two types of shape changes (holding color and size of objects constant) – a within-kind shape change (e.g. a sippy cup with two handles vs. a regular cup with one handle) or a cross-kind shape change (e.g. a cup and a bottle). During habituation trials, we found that the infants were equally sensitive to both types of shape change. On the test trials, however, only the infants who saw the cross-kind shape change showed evidence of successful object individuation by looking longer at the one-object, unexpected outcome than the two-object, expected outcome. These results provide preliminary evidence that kind representations (and not just property representations) underlie the success at 12 months.

Furthermore, the capacity to individuate in the absence of spatiotemporal information that emerges between 10 and 12 months of age is closely tied to linguistic competence in ways that implicate kind concepts. Xu and Carey (1996) found that 10-month-olds who knew the labels for the objects succeeded at individuating familiar objects on the basis of kind distinctions (the objects were a ball, a bottle, a book, and a cup). A new set of studies showed that labeling facilitates individuation in this paradigm. Xu (1998) tested 9-month-old infants using the Xu and Carey (1996) paradigm and gave the infants verbal labels for the objects. When the toy duck emerged from behind the screen, the experimenter said, in infant directed speech, “Look, [baby’s name], a duck”. When the duck returned behind the screen and the ball emerged from the other side, the experimenter said, “Look, [baby’s name], a ball”. On the test trials, infants were shown an expected outcome of two objects, a duck and a ball, or an unexpected outcome of just one object, a duck or a ball. Infants looked longer at the unexpected outcome of a single object. In a control condition, infants heard “a toy” for both the duck and the ball, and their looking time pattern on the test trials was not different from their baseline preference. In a second study, two tones were used instead of two labels and infants again failed to look longer at the one-object outcome. Our interpretation of this finding is that contrasting labels provide signals to the infant that two kinds of objects are present, and that there must therefore be two numerically distinct objects behind the screen. The negative finding with tones suggests that perhaps language in the form of labeling plays a specific role in signaling object kinds for the infants. It is unclear whether labels are necessary for the formation of kind representations (cf. the experiments of Mandler and her colleagues cited below; we are agnostic as to the format of representation of symbols for kinds). We take these results as part of a general pattern of findings that infants expect labels to refer to kinds, and that kind membership has consequences for both individuation and categorization (e.g. Balaban & Waxman, 1997; Waxman, 1999).

Kind concepts differ from property concepts in ways other than that kinds provide criteria for individuation and numerical identity. Other infant studies confirm that kind representations are differentiated from property representations by the end of the first year of life (see Mandler, 2000; Xu & Carey, 2000, for reviews), and that labeling facilitates kind representations (Balaban & Waxman, 1997; Waxman & Markow, 1995). Furthermore, Waxman (1999) showed that by 13 months, infants distinguish linguistically between kind representations and property representations. Upon hearing a series of objects described by a count noun (“Look, its a blicket”) they extract kind similarity (at both the basic and superordinate levels) but not property similarity (texture and color), whereas upon hearing an adjective (“Look, its a blickish one”) they extract property similarity as well.

In sum, these studies support the claims that kind representations are architecturally distinct from property representations, as they play distinct roles in individuation, categorization, and language. These studies also lend support to the architectural distinction between object file-based individuation and kind-based individuation, for this latter system emerges markedly later in development.

6.2. Lessons concerning the mid-level object file/object tracking system itself

Suppose it is true that kind-based individuation is architecturally distinct from the mid-level object tracking system, that the mid-level system underlies object individuation and tracking early in early infancy and that the kind-based system is not developed until the end of the first year of life. If so, studies of young infants provide us with a wonderful methodological tool – a chance to study the object tracking system pure, so to speak, uncontaminated by kind representations. Before the emergence of the kind-based system, the processes that create representations of individual objects create only object files. Properties of objects are represented as features bound to object files. After this developmental change, the processes that create representations of individual objects also create symbols for kind sortals, such as *duck*, and properties of these individuals may be bound directly to them, as in *yellow duck*. Once this second system of kind-based object individuation has become available, it creates the representations that articulate thought. That is, it preempts object file representations in our experiences of the world. This is why, in the absence of direct spatiotemporal evidence to the contrary, we infer that the duck and the cat moved in Fig. 1, and why we consider that a person ceases to exist when he or she dies, in spite of the spatiotemporal continuity of bodies. Thus, for adults, we need to set up situations that prevent the operation of the second system (high attentional load, as in MOT or search studies, or very brief exposures, as in apparent-motion studies or feature conjunction studies) or situations that separate perception from judgment in order to study the operation of the mid-level object file and object tracking systems.

If we accept the arguments of the paper so far, then the study of object representations in very young infants can provide invaluable evidence concerning the nature of the mid-level systems, for very young infants do not yet have available the kind-based systems which preempt the output of the mid-level vision in adult conceptual representations. In the remaining sections of this paper, we sketch what might be learned about the object file and object tracking system from studies of the object representations of very young infants.

6.3. Short-term memory and object file representations

In MOT experiments and in studies of object-based attention in which the objects undergo real or apparent motion (Kahneman et al., 1992), subjects are in nearly continuous visual contact with the objects. Occlusion, if present at all, is momentary. FINSTs are indices that depend upon spatiotemporal information in order to remain assigned to individuals. It is unclear from these studies, then, whether object files are stable object representations that may be placed into longer lasting short-term memory stores, perhaps even losing their current spatiotemporal indices. The object permanence and number studies of young infants suggest that they can.

Many of the infant studies cited above involve occlusion, sometimes for as long as 10 s or more. In Wynn's $1 + 1$ (or $2 - 1$) studies, for instance, the first object (or pair) remains hidden for several seconds, and a memory representation of that object (or pair) must be updated, in memory, as the result of the addition (or subtraction). Then when the outcome array is revealed object file representations are again computed, and the resultant models (the short-term memory object file representation, and the current outcome object file representation) are compared.[8] Koechlin et al. (1998) showed that 5-month-old infants succeed in $1 + 1 = 2$ or 1 and $2 - 1 = 2$ or 1 addition/subtraction studies even when the objects behind the screen are placed on a rotating plate. Under these conditions, the infant cannot maintain an index on a hidden object; that is, when the outcome is revealed, the infant has no way of knowing which object on the plate is the same object as the first one placed behind the screen and which one is the same object as the second one. This finding supports the assumption by Simon (1997) and Uller et al. (1999) that *two* object file models, one of the set-up event and one of the outcome array, are being compared.

Feigenson, Carey, and Hauser (2001) have new findings that lend support to the hypothesis that the infant can create and store more than one memory model of sets of objects, and compare them numerically in memory. Furthermore, these studies show that the total number of objects represented in two separate short-term memory stores can exceed the limits of object indexing, showing that short-term memory stores may include object files that are not currently indexed. Ten-month-old infants were shown a given number of graham crackers placed into one box, and a different number placed into another box. The infants could not see the crackers in the box. The infants watched the crackers being placed into the two boxes, and then they were allowed to crawl to one or the other. At issue was whether they would go to the box with the larger number of crackers. This is what they did, when the choice was 1 vs. 2 or 2 vs. 3. Performance fell apart at 3 vs. 4 and at 3 vs. 6. This latter finding is important, for it rules out that analog magnitude number representations (see Dehaene, 1997; Gallistel, 1990, for characterizations of and evidence for analog magnitude number representations of number) could be underlying performance on this task. Success when analog magnitude number representations are activated is a function of the ratio between the set sizes; 3 vs. 6 is the same Weber fraction as 1 vs. 2 and is more discriminable than 2 vs. 3. Success within the range of parallel individuation and failure outside it, controlling for ratio, is the set size signature of object file representations of these individuals. This is the earliest demonstration of an ordinal quantitative judgment in infants, but it is the success at 2 vs. 3 that is of theoretical importance in the present context. Sets of 2 or 3 objects are each within the infants' limits of object indexing, but sets of 5 are not. Thus, infants cannot be indexing a single set of objects in this experiment. Rather, they must be establishing two short-term memory models, one consisting of 2 object files and one consisting of 3 object

[8] In the Simon (1997) and Uller et al. (1999) accounts of these experiments it was assumed that the comparisons were based on $1 - 1$ correspondence among object files in the two models. Subsequent experiments (Clearfield & Mix, 1999; Feigenson et al., 2001, in press) make it clear that object file models are often compared on the basis of total surface area or volume, or on the basis of properties of the individual objects, and that these properties of object file representations are more salient than is the number of object files in a model. These facts do not undermine the conclusion that object file representations are underlying the infants' behavior in these studies, but they do undermine the conclusion that these experiments reflect numerical computations over object file representations.

files, and then comparing them in memory.[9] Thus, object file representations do not merely underlie momentary tracking of objects. Rather, object files are symbols that articulate relatively long lasting short-term memory models, which, in turn, support other computations; in this case, comparisons with respect to more or less.

6.4. Mid-level object representations: preconceptual? Or, what kinds of things are FINGs?

Pylyshyn (2001) suggests that the individuals that are indexed in the mid-level object tracking experiments are non-conceptual. Of course, the individuals that are indexed are in the world (hence neither preconceptual or conceptual). At issue is whether the symbols for these individuals, the object files themselves, are preconceptual or conceptual symbols. Recall that Pylyshyn (2001) agrees that the assignment of a FINST is the initial phase of creating an object file, and thus that FINGs (the individuals FINSTs point to) are the same individuals as those represented by object files.

We have discussed at length one sense in which object files are preconceptual symbols; they do not represent object kinds such as *dog* or *cup*. In addition, Pylyshyn (2001) is mainly concerned with the issue of whether the processes that use features or spatiotemporal information to assign indexes are themselves conceptual processes. He argues that individuals are picked out by perceptual processes, perhaps in a bottom-up manner; individuals are *not* determined by a process that examines explicitly represented definitional or probabilistic features, even spatiotemporal ones.

Although we believe that Pylyshyn is right about this, the question still remains concerning object files as symbols themselves. Notice that the fact that perceptual processes (figure/ground segregation, surface representation, object tracking on the basis of spatiotemporal information) establish object files does not make them perceptual symbols. Perceptual processes may deliver symbols that are conceptual, as seen by their conceptual role.

An analogy may clarify our argument here. Michotte (1963) specified the spatiotemporal parameters of the relation between two moving bodies sufficient for the perception of causal interaction, e.g. for the perception that contact with one moving body caused a second one to move. That there are perceptual processes that yield representations of causality does not mean that that these representations themselves are perceptual. Causal attribution transcends the spatiotemporal parameters, being contributed by the mind, and guides further inferences and actions, being in that sense informationally promiscuous. In these senses, then, representations of caus-

[9] See previous footnote. In the infant choice experiment, infants were maximizing the total amount of graham cracker. Given a choice between one large cracker in one container and two small crackers, summing to half the volume of the large one, infants chose the single large cracker. Still, the set size signature of object file representations obtained success at 2 vs. 3, but not at 3 vs. 6, indicating that the comparison was mediated by object file representations and not representations that could keep a running total of volume apart from the individual objects.

ality are conceptual, even though there are dedicated perceptual processors that compute them.

To explore the issue of whether object files are conceptual symbols, we must begin by considering their *content*. What do object files represent? Two types of empirical evidence bear on this question: (1) studies of the extensions of object files (What entities in the world cause object files to be established? What are FINGs (an empirical question)?) and (2) studies of the conceptual role of object files (What computations do object file symbols participate in?). We shall argue that the content object files is *physical objects*, by which we mean what is sometimes called "Spelke-objects", namely, bounded, coherent, 3D, separable and moveable wholes. And we will argue that object file representations are conceptual in the sense that they articulate physical reasoning, enter into number-relevant computations, and support intentional action. Sections 6.5–6.7 review the evidence in support of these claims.

6.5. The extension of object files

The claim that object files represent real 3D objects may seem hardly surprising, but in fact, there are reasons to doubt it. The arrays are *actually* 2D objects in virtually all of the adult studies on mid-level vision, as well as in some of the infant studies (e.g. those of Johnson, 2000, and his colleagues on amodal completion behind barriers and those of Johnson and Gilmore, 1998, on object-based attention). But because we can present many of the cues for depth in 2D arrays, surfaces arrayed in 3D are routinely perceived in such displays. That the system can be fooled (similarly for Michotte causality) does not mean that it is not representing the stimuli as Spelke-objects. What reasons do we have for thinking that this may be the case?

We have already presented one line of evidence that object files represent Spelke-objects. The processes that establish and maintain object file representations are sensitive to the distinction between the spatiotemporal information that specifies occlusion, on the one hand, and that that specifies the cessation of existence, on the other. Occlusion and existence cessation are properties of real physical objects. Furthermore, studies of infants shown pictures suggest that infants sometimes misperceive 2D representations as if they were real 3D objects. Many studies have shown that infants attempt to grasp pictured objects well into the second year of life (see DeLoache, Pierroutsakos, Uttal, Rosengren, & Gottlieb, 1998).

Two series of studies with 8-month-old infants underline the point that the individuals being tracked in the infant studies are physical objects, and not just any perceptual objects specified by figure/ground processes. A hallmark of physical objects is that they maintain their boundaries through time. Neither a pile of sand nor a pile of blocks is a Spelke-object, in spite of the fact that when stationary it may be perceptually indistinguishable from one. One may make a pile-shaped cone and coat it with sand, or one may put together a set of small objects, yielding a single pile-shaped entity. It is only upon viewing such entities in motion (do they fall apart, or do they maintain their boundaries?) that unequivocal evidence for their ontological status is obtained. Infants track Spelke-objects that are perceptually identical to piles of sand (Huntley-Fenner, Carey, & Salimando, 2001) or piles of little blocks (Chiang & Wynn, in press) under conditions where they will not track the perceptually identical non-objects.

Take Huntley-Fenner et al. (2001) for example. They carried out $1 + 1 = 2$ or 1 studies involving sand poured behind or sand objects being lowered behind the barriers. When the sand was resting on the stage, it formed a pile, and the sand objects, when resting on the stage, were perceptually indistinguishable, being pile-shaped objects coated in sand. It was only upon seeing the entity being poured (sand) or lowered (object) onto the stage that infants could identify the resulting pile-shaped entity as sand or as an object. Stimulus type was a between-participant variable, and infants were familiarized with the stimuli before the study by handling the sand or the sand object. One study involved a single screen; another involved two screens. Eight-month-old infants succeeded in the sand object conditions, but failed in the sand conditions. The failure in the two-screen study is especially striking, for it shows that infants do not have "sand permanence". In this study, the infant watched as a pile of sand was poured onto the stage floor, and then hidden behind a screen. A second, spatially separate, screen was introduced and a second pile of sand was poured behind it. The screens were then removed, revealing either two piles of sand (one behind each screen) or only one (the original pile seen on the stage floor initially). Eight-month-olds did not differentiate the two outcomes, although they succeeded if the stimuli were sand-pile-shaped Spelke-objects lowered as a whole onto the stage floor. As mentioned above, object permanence requires an individual whose identity is being tracked; it is the *same individual* we represent behind the screen. Apparently, 8-month-old infants cannot establish representations of individual portions of sand and trace them through time.

These infant studies suggest that the object tracking system is just that: an *object* tracking system, where *object* means 3D, bounded, coherent physical object. It fails to track perceptually specified figures that have a history of non-cohesion. That the system can be fooled, can *misrepresent* 2D stimuli as objects, does not militate against this conclusion.

One final line of work on infant object representations bolsters this conclusion. Identical spatiotemporal principles (e.g. independent motion) specify tactile and visual objects, and infants map representations across the two modalities. Streri and Spelke (1988) allowed young infants to handle rings (one in each hand) that they could not see. When the rings moved independently of each other, infants preferred to look at a display containing two spatially separate objects. In contrast, when handling rings connected by a rigid rod (again, one in each hand), such that they did not move independently of each other, they preferred to look at a display containing a single object. (In cross-modal experiments of this sort, infants typically prefer to look at the visual stimulus that matches the tactually represented stimulus, presumably because they seek a consistent representation of their world.)

In sum, infant object representations appear to have 3D, bounded, coherent, separately moving objects in their extensions. On the assumption that infant object representations are object files, we conclude that "object files" are well named: they represent real physical objects.

6.6. Conceptual role: object file representations are the input into volitional action

Section 6.5 concerns what real world individuals are represented by object files. This is one part of the project of specifying the content of a symbol; the other part is specifying its conceptual role. Files representing currently visible attended objects, as well as those stored in short-term memory, guide actions directed towards the physical world. By 8 months, infants solve Stage 4 object permanence tasks (retrieving objects hidden under cloths, behind barriers). Similarly, at 10 months, before kind representations support individuation, object file representations support manual search in the Van de Walle et al. (in press) object retrieval tasks and in the Feigenson et al. (2001) number comparison experiments cited above. Insofar as being available to guide volitional action (informational promiscuity) is evidence that a representation is conceptual, these studies suggest that object files are.

6.7. Conceptual role: object representations articulate physical knowledge

The actions in the Feigenson et al. studies were based on the output of computations that established which container contained more crackers. That object file representations enter into comparative quantity computations suggests that they have conceptual roles that far transcend merely representing objects that the infant may reach for. Indeed, it is in the exploration of the conceptual role of object representations that the infant studies most dramatically transcend the literature on mid-level vision, for these studies have not been concerned with the inferences that are drawn about objects. If the identification of the infant's object representations with object files is correct, then these studies show that object file representations articulate considerable physical knowledge. Some of this physical knowledge may be innate, instantiated in the computations that establish representations of object files in the first place. But other aspects are learned – object files are representations of objects about which infants can learn, and in this learning they learn about objects as a class, not just about individual object tokens.

6.7.1. Innate physical knowledge about objects

By 2 months of age, infant object file representations are quite adult-like. For example, Johnson, 2000 reviews the literature on surface perception in infancy. By 2 months of age, infants are sensitive to almost all the same information adults are in building representations of the amodally complete surfaces behind barriers, although young infants need more redundant cues than do older children or adults. Astoundingly, 2-month-olds are also able to represent physical relations such as inside and behind, and their representations are constrained by knowledge of solidity, a property of Spelke-objects but not of 2D visual objects. Spelke et al. (1992) habituated 2-month-olds to a ball rolling behind a screen, the screen then being removed and the ball shown resting against the back wall. They then inserted a barrier behind the screen, perpendicular to it with its top visible, and rolled the ball behind again. Upon removal of the screen, infants looked longer if the ball ended up against the back wall, having apparently passed through the solid barrier, than if the ball was revealed resting against the barrier. Convergent evidence is provided by Hespos and Baillargeon (in press), who showed that 2-month-olds expect objects inside other objects to move with them, in contrast to objects behind other objects, and also that they expect objects can be inserted into open containers but not into closed containers (the latter being a violation of solidity).

Besides expecting objects to be solid, and thus not to pass through other ones, by 6 months infants also expect objects to be subject to the laws of contact causality (Leslie & Keeble, 1986). Young infants look longer if an object goes into motion without having been contacted by another moving object than if it has (Spelke, Philips, & Woodward, 1995) and they look longer if a small object hitting another makes it move farther than if a larger object going the same speed does (Baillargeon, 1995).

Thus, the conceptual role of the infant's object representations is that of 3D Spelke-objects; objects are represented as solid entities in spatial relations with each other that cannot pass through other objects, and which move only upon contact. If we accept the identification of the infant's object concept with object files, then we must accept that object file representations also have the same conceptual role.

6.7.2. Learning generalizations about objects

Still under debate is what aspects of the conceptual role of object representations described above are innate and what are learned. There is no doubt, however, that infants learn many generalizations about objects during their early months. Thus, the processes that yield object representations yield representations about which the infant learns. To take just one example – infants do not innately know that unsupported objects fall (Baillargeon, 1995). That is, if they watch an object slowly pushed off a platform until it is completely unconnected to it, apparently suspended in mid-air, 3-month-olds show no differential interest relative to whether it is adequately supported from below. Just a few weeks later, though, this event draws long looking, relative to events in which the object is supported. In a series of beautiful experiments, Baillargeon has shown that infants' learning about support unfolds in a regular way. First they are not surprised that the object does not fall so long as there is any contact with the support, then the contact must be from below, then more than half of the base of the object must be supported from below, and finally they take into account the geometry of the object. Furthermore, the initial stages of this learning occur, in the ordinary course of events, from infants' own attempts to place objects on surfaces, but it can also be driven from observational evidence alone.

One important conclusion from these studies is that they reveal generalizations that infants make about *objects*; experience placing stuffed animals on tables enables infants to predict whether any unsupported Spelke-object will fall. Systematic study of generalization from observational evidence would be of great interest in constraining our models of the learning process. At the very least, infants have not had previous experience with the specific objects in the Baillargeon support studies. That is, physical reasoning about Spelke-objects embodies knowledge formulated over the category *object*, whatever the format of this knowledge.

6.8. Interim conclusions: what are object files symbols of?

Two lines of evidence support the conclusion that infants' object representations have Spelke-objects as their content. First, the extensions of the symbols seem to be real 3D, bounded, coherent objects. Infants do not track individuals that cannot be construed as Spelke-objects, like piles of sand or piles of blocks, or entities that shrink to nothing or explode. Infants sometimes attempt to pick up pictured objects, providing evidence that they sometimes misconstrue 2D representations of objects as Spelke-objects. And infants have cross-modal representations of individuated 3D objects; not only do the same principles specify object number, but infants map the object representations built on tactile spatiotemporal evidence to visual representations of objects. Second, studies of the conceptual role of object representations show that they support action, quantitative comparisons, and articulate physical reasoning. If we accept the identification of infants' object representations with object files, then we must correspondingly enrich our conception of the latter.

7. A summary overview

This paper is speculative. We do not know for sure that young infants' object representations are identical to those computed by the mid-level object-based attention system. As one reviewer pointed out, it may be that the two are quite distinct representational systems, and their similarities reflect the fact that both are designed to solve similar problems – picking out individuals and tracking them through time. Of course this is possible, but we doubt it, for the similarities we draw upon in making the identification are non-veridical. Objects do not change color and texture over the short time course in which both systems allow object representations to be updated, and there is no particular reason for the limitations on the set size of objects that may be individuated in parallel to be so similar if the systems are distinct. But these are early days in exploring the relations between the two literatures, and no doubt in many details our speculations will turn out to be wrong.

We have argued here that the discovery, if true, that young infants' object representations are the same natural kind as the object files of mid-level vision has important consequences for both literatures. Merging the two literatures brings new data to bear on very general theoretical disputes within each literature, such as the content of object representations, the relative roles of spatiotemporal, featural and kind information in object individuation and tracking, and the senses in which object representations are preconceptual and the senses in which they are conceptual.

Acknowledgements

The research reported here was supported by NSF grants (SBR-9712103 and SBR-951465) to S.C., and NIH B/START grant (R03MH59040-01) and NSF grant (SBR-9910729) to F.X. We thank Zenon Pylyshyn, Brian Scholl, Cristina

Sorrentino, Joshua Tenenbaum, Gretchen Van de Walle, and two anonymous reviewers for helpful discussion and very helpful comments on an earlier draft.

References

Aguilar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology*, *39*, 116–157.

Baillargeon, R. (1991). Reasoning about the height and location of a hidden object in 4.5- and 6.5-month-old infants. *Cognition*, *38*, 13–42.

Baillargeon, R. (1995). A model of physical reasoning in infancy. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9, pp. 305–371). Norwood, NJ: Ablex.

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: further evidence. *Child Development*, *62*, 1227–1246.

Baillargeon, R., Miller, K., & Constantino, J. (1993). *Ten-month-old infants' intuitions about addition*. Unpublished manuscript, University of Illinois at Urbana, Champaign, IL.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*, 3–26.

Bower, T. G. R. (1974). *Development of infancy*. San Francisco, CA: W.H. Freeman.

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, *4*, 121–138.

Carey, S., & Bassin, S. (1998). *When adults fail to see the trick. Adult judgments of events in an infant violation of expectancy looking time study*. Poster presented at the 11th biennial meeting of the International Society for Infant Studies, Atlanta, GA.

Carey, S., & Xu, F. (1999). Sortals and kinds: an appreciation of John Macnamara. In R. Jackendoff, P. Bloom, & K. Wynn (Eds.), *Language, logic, and concepts: essays in honor of John Macnamara*. Cambridge, MA: MIT Press.

Chiang, W. C., & Wynn, K. (in press). Infants representations and teaching of objects: implications from collections. *Cognition*.

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10* (5), 408–411.

Dehaene, S. (1997). *The number sense: how the mind creates mathematics*. Oxford: Oxford University Press.

DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science*, *9* (3), 205–210.

Feigenson, L., Carey, S., & Hauser, M. (2001). Infants' spontaneous ordinal choices, submitted for publication.

Feigenson, L., Carey, S., & Spelke, E. S. (in press). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gupta, A. (1980). *The logic of common nouns*. New Haven, CT: Yale University Press.

Hespos, S., & Baillargeon, R. (in press). Knowledge about containment events in very young infants. *Cognition*.

Hirsch, E. (1982). *The concept of identity*. New York: Oxford University Press.

Huntley-Fenner, G., Carey, S., & Salimando, A. (2001). Objects are individuals but stuff doesn't count: perceived rigidity and cohesiveness influence infants' representation of small numbers of discrete entities, submitted for publication.

Johnson, M. H., & Gilmore, R. O. (1998). Object-centered attention in 8-month-old infants. *Developmental Science*, *1* (2), 221–225.

Johnson, S. (2000). The development of visual surface perception: insights into the ontogeny of knowledge. In C. Rovee-Collier, L. Lipsitt, & H. Hayne (Eds.), *Progress in infancy research* (Vol. 1, pp. 113–154). Mahwah, NJ: Erlbaum.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: object specific integration of information. *Cognitive Psychology*, *24*, 175–219.

Koechlin, E., Dehaene, S., & Mehler, J. (1998). Numerical transformations in five-month-old infants. *Mathematical Cognition*, *3*, 89–104.

Leslie, A. M., & Keeble, S. (1986). Do six-month-old infants perceive causality? *Cognition*, *25*, 265–288.

Leslie, A., Xu, F., Tremoulet, P., & Scholl, B. (1998). Indexing and the object concept: developing "what" and "where" systems. *Trends in Cognitive Sciences*, *2* (1), 10–18.

Macnamara, J. (1986). *A border dispute: the place of logic in psychology*. Cambridge, MA: MIT Press.

Mandler, J. M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development*, *1*, 3–36.

Marr, D. (1982). *Vision*. New York: Freedman.

Michotte, A. (1963). *The perception of causality*. New York: Basic Books.

Michotte, A., & Burke, L. (1951). *Une novelle enigme de la psychologie de la perception: le "donee amodal" dans l'experience sensorielle*. Aces du 13 eme Congrages Internationale de Psychologie, Stockholm, pp. 179–180.

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: a critical link between lower-level and higher-level vision. In S. M. Kosslyn, & D. N. Osherson (Eds.), *Visual cognition* (2nd ed., pp. 1–70). Cambridge, MA: MIT Press.

Needham, A. (1998). Infants' use of featural information in the segregation of stationary objects. *Infant Behavior and Development*, *21* (1), 47–76.

Needham, A., & Baillargeon, R. (1997). Object segregation in 8-month-old infants. *Cognition*, *62*, 121–149.

Needham, A., & Baillargeon, R. (1998). Effects of prior experience on 4.5-month-old infants' object segregation. *Infant Behavior and Development*, *21* (1), 1–24.

Needham, A., & Baillargeon, R. (2000). Infants' use of featural and experiential information in segregating and individuating objects: a reply to Xu, Carey, & Welch (1999). *Cognition*, *74*, 255–284.

Needham, A., Baillargeon, R., & Kauffman, L. (1997). Object segregation in infancy. In C. Rovee-Collier, & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 11, pp. 1–39). Greenwich, CT: Ablex.

Peterson, M. A. (1994). Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science*, *3* (4), 105–111.

Peterson, M. A., de Gelder, B., Rapcsak, S. Z., Gerhardstein, P., & Bachoud-Levis, A. (in press). A double dissociation between conscious and unconscious object recognition processes revealed by figure-ground segregation. *Vision Research*.

Peterson, M. A., & Gibson, B. (1993). Shape recognition inputs to figure-ground organization in three-dimensional displays. *Cognitive Psychology*, *25*, 383–429.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects and situated vision. *Cognition*, this issue, *80*, 127–158.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, this issue, *80*, 1–46.

Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept: beyond the perception/cognition dichotomy. In E. Lepore, & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 26–73). Oxford: Blackwell.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: clues to visual objecthood. *Cognitive Psychology*, *38*, 259–290.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multi-element tracking. *Cognition*, this issue, *80*, 159–177.

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (2001). The relationship between property-encoding and object-based attention: evidence from multiple-object tracking, submitted for publication.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: a "non-numerical" account. *Cognitive Development*, *12*, 349–372.

Simon, T., Hespos, S., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, *10*, 253–269.

Slater, A., Johnson, S. P., Brown, E., & Badenoch, M. (1996). Newborn infants' perception of partly occluded objects. *Infant Behavior and Development*, *19*, 145–148.

Spelke, E. S. (1996). Initial knowledge: six suggestions. *Cognition*, *50*, 431–445.

Spelke, E. S., Brelinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*, 605–632.

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatio-temporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, *13*, 113–142.

Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: a multidisciplinary debate*. Oxford: Clarendon Press.

Starkey, P., & Cooper, R. (1980). Perception of numbers by human infants. *Science*, *210* (28), 1033–1034.

Streri, A., & Spelke, E. S. (1988). Haptic perception of objects in infancy. *Cognitive Psychology*, *20*, 1–23.

Trick, L., & Pylyshyn, Z. (1994). Why are small and large numbers enumerated differently? A limited capacity preattentive stage in vision. *Psychological Review*, *101*, 80–102.

Uller, C., Huntley-Fenner, G., Carey, S., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development*, *14*, 1–36.

Uller, C., & Leslie, A. (1999, April). *Assessing the infant counting limit*. Poster presented at the biennial meeting of the Society for Research on Child Development, Albuquerque, NM.

Van de Walle, G., Carey, S., & Prevor, M. (in press). The use of kind distinctions for object individuation: evidence from reaching. *Journal of Cognition and Development*.

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition*, *70*, B35–B50.

Waxman, S. R., & Markow, D. R. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.

Wiggins, D. (1980). *Sameness and substance*. Oxford: Basil Blackwell.

Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition*, *72*, 125–166.

Wilcox, T., & Baillargeon, R. (1998a). Object individuation in infancy: the use of featural information in reasoning about occlusion events. *Cognitive Psychology*, *37*, 97–155.

Wilcox, T., & Baillargeon, R. (1998b). Object individuation in young infants: further evidence with an event-monitoring paradigm. *Developmental Science*, *1*, 127–142.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749–750.

Xu, F. (1997). From Lot's wife to a pillar of salt: evidence that physical object is a sortal concept. *Mind and Language*, *12*, 365–392.

Xu, F. (1998). Distinct labels provide pointers to distinct sortals in 9-month-old infants. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development* (pp. 791–796). Somerville, MA: Cascadilla Press.

Xu, F. (1999). Object individuation and object identity in infancy: the role of spatiotemporal information, object property information, and language. *Acta Psychologica*, *102*, 113–136.

Xu, F. (2000). *Numerical competence in infancy: two systems of representations*. Paper presented at the 12th International Conference on Infant Studies, Brighton.

Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. *Cognitive Psychology*, *30*, 111–153.

Xu, F., & Carey, S. (2000). The emergence of kind concepts: a rejoinder to Needham & Baillargeon. *Cognition*, *74*, 285–301.

Xu, F., Carey, S., Quint, N., & Bassin, S. (2001). Kind-based object individuation in infancy. Manuscript in preparation.

Xu, F., Carey, S., & Welch, J. (1999). Infants' ability to use object kind information for object individuation. *Cognition*, *70*, 137–166.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.