COGNITION

Cognition 74 (2000) 285–301

www.elsevier.com/locate/cognit

Discussion

The emergence of kind concepts: a rejoinder to Needham and Baillargeon (2000)

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

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

1. Introduction

Perhaps Needham and Baillargeon would be surprised to find that we are in complete agreement on many important points, including: (1) at issue is the nature of infants' representations of events and objects, and how infants parse the world into individuals and kinds; (2) looking time methodologies are a fruitful source of data that bear on these issues; (3) under some circumstances, infants below 12 months of age make use of experiential, physical (including spatiotemporal) and featural information for object individuation; (4) it is of utmost importance to resolve apparent conflicts in data and interpretations from different studies; (5) finally, our results do seem to conflict with results from experiments by Needham, Wilcox, and Baillargeon, and relevant data from other laboratories are rather mixed (see below).

In our experiments (Xu & Carey, 1996; Xu, Carey & Welch, 1999), 10-month-old infants failed to use property/featural information (e.g. the differences in shape, color, or texture) or object kind information (e.g. the difference between a duck and a car) to parse a display into two objects (although under some circumstances, success is obtained at 9 months; see Section 5.1). By 12 months of age, infants succeeded at these tasks. In contrast, Needham and Baillargeon (henceforth N&B) reviewed studies which provide evidence that under some circumstances, infants succeeded in using property/featural information for object individuation as early as 4.5 months. How do we resolve this apparent conflict?

In their reply, N&B questioned both our results and our interpretation of the results. The results, however, are highly robust – the studies of Xu and Carey (1996) had four internal replications of the failure at 10 months, and the shift

* Corresponding author. Fax: + 1-617-373-8714.
E-mail address: fxu@neu.edu (F. Xu)

PII: S0010-0277(99)00076-1

between 10 and 12 months has been replicated by Wilcox and Baillargeon (1998a). Similarly, Xu et al. (1999) had three internal replications of the failure at 10 months. (See also Sections 2.1 and 3.2 below for additional evidence.)

Our working hypothesis is that one major source of the change from 10 to 12 months is the emergence of genuine kind concepts, e.g. *duck*, *car*, *cup*. Furthermore, we suggest that our tasks require access to symbols for kinds, whereas the tasks of Needham, Wilcox, and Baillargeon do not.

N&B doubted our working hypothesis on several grounds. First, they pointed out, correctly, that we have not provided evidence that success at 12 months was based on kind representations rather than property representations. Second, they suggested that some of our findings may be artifactual, and more generally, they suggested that the information-processing demands of our tasks, which are greater than those of Needham, Wilcox, and Baillargeon, may have masked the infant's competence at using property or kind information for object individuation. Lastly, they provided evidence that infants as young as 4.5 months (the box/cylinder studies; Needham, 1998) and 7.5 months (the key-ring studies and the three-box studies; Needham, 1999) used experiential knowledge for object individuation. The use of experiential knowledge is taken as evidence that infants younger than 10–12 months already represent object kinds, such as box or key-ring.

In our rejoinder, we will answer the specific criticisms raised by N&B, and offer an alternative resolution of the conflicting results between the laboratories, one that is supported by additional empirical evidence.

2. Replies to specific criticisms

2.1. Property or kind representations at 12 months?

One very important theoretical and interpretative issue raised by N&B concerns the basis of success at 12 months in our studies. One possibility is that infants' success was based on kind representations – infants are drawing on kind distinctions, such as that between the kind duck and the kind ball. The other possibility is that success was based on property representations – infants are drawing on distinctions, such as that between irregular-shaped, yellow/orange, and rubbery, on the one hand, and round, white, and shiny, on the other. We explicitly draw attention to this interpretative ambiguity in our papers (Xu & Carey, 1996; Xu et al., 1999). In fact, we labeled the experimental condition 'property/kind condition' for precisely this reason.

To address this problem, we have recently completed a series of experiments with 12-month-olds (Xu, Carey & Quint, 1997). 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 returned. 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 versus 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.

2.2. Idiosyncratic factors determine patterns of success and failure

N&B pointed out, rightly, that factors that are idiosyncratic to any given experimental paradigm contribute to task success. One nice example they offered, and provided preliminary evidence for, is that stacked objects that share a boundary are more difficult to individuate than are the same two objects side by side (Needham, 1999). Although the source of the difficulty is still unclear, the fact that the objects in Xu et al. (1999) were stacked might have contributed to the information-processing demands of this task.

N&B also pointed out that the hand changed its action on the test trials of the static conditions, but not the movement conditions, of Xu et al. (1999). They suggested that the 10-month-olds in the static conditions were therefore drawn to the movement of the hand on the test trials, and this masked their correct parse of the display as a duck and a car. The change between 10 and 12 months, on this alternative, reflects the older infants being less distracted by the hand, rather than their being better able to individuate the objects on the basis of property or kind differences. This is an empirical possibility (though we doubt it is right), on which there are as of now no relevant data. The data N&B offer as consistent with their interpretation (that 10-month-olds are equally interested in both outcomes on the test trials in the static condition, similar to the level of interest in the impossible outcome in the movement condition) is equally consistent with our interpretation. Our account of this detail of the data is as follows: precisely because the 10-month-old infants were unsure about how to parse the duck–car (or the cup–shoe) display, both test outcomes were interesting to them. That is, both outcomes gave them a definite parse of the display and if the infant was in a state of uncertainty, she may well attend to both of them.

2.3. The 10- to 12-month shift – a massive coincidence?

One way to interpret N&B's reply to our work is that they may be suggesting that different idiosyncratic factors in each of our studies differentiate the performance of the 10-month-olds from that of the 12-month-olds. This is, of course, always an empirical possibility, and underlies our search for further convergent evidence with different methodologies (see Section 3.2 below). However, we note that there is abundant evidence in the literature for failures of young infants to draw on property/featural information for object individuation, in sharp contrast to successes at using spatiotemporal evidence under nearly identical experimental circumstances, using both looking time and simple reaching measures (Johnson & Aslin, 1996; Jusczyk, Johnson, Spelke & Kennedy, 1999; Kellman & Spelke, 1983; Kellman, Spelke & Short, 1986; Kellman, Gleitman & Spelke, 1987; Simon, Hespos & Rochat, 1995; Spelke, 1990; Spelke, Breinlinger, Jacobson & Phillips, 1993; Spelke, Kestenbaum, Simons & Wein, 1995; Streri & Spelke, 1988; von Hofsten & Spelke, 1985; see also, studies with infants using food items which provide convergent evidence, Spelke, pers. commun.). We note that in many of these studies the objects are side by side rather than stacked, and when stacked the hand touches the top object both during habituation and test. Given the systematicity of this pattern of findings, from other laboratories as well as our own, we very much doubt the 'massive coincidence' account of our data.

3. A major component of N&B's account: information-processing demands

Less idiosyncratically, N&B pointed out, again rightly, that the information-processing demands of our tasks are greater than theirs. In particular, they draw attention to three aspects of the tasks with consequences for information processing: object complexity, the distinction between event-monitoring and event-mapping, and the difference between single trajectory versus multiple emergences. We admire and applaud the very elegant experiments by Wilcox and Baillargeon demonstrating the influence of these factors on task success. It is important, however, to probe deeper into the reasons for these effects. We begin by questioning whether the success in the narrow-screen experiments reflects the infants' use of property information for object individuation.

3.1. A reinterpretation of the narrow-screen/wide-screen results

The narrow-screen studies of Wilcox and Baillargeon (1998a,b) (henceforth W&B) were designed to test the hypothesis that infants would succeed earlier in an event monitoring paradigm than in the traditional event-mapping paradigms. In these 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. W&B contrasted two conditions. 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 only 21 cm wide, too narrow for both objects to fit behind. W&B found that infants as young as 4.5 months of age looked longer at the narrow-screen event than the wide-screen event. By the logic of violation of expectancy looking time measures (a logic we accept), infants apparently found the narrow-screen event anomalous, relative to the wide-screen event. W&B interpreted the infants' behavior as follows: in the narrow-screen event, the infants must have used the property (or kind) differences between the box and the ball to infer two distinct objects and realized that the two objects could not fit behind the screen simultaneously. Furthermore, Wilcox (1999) found a progression in the ages at which infants succeed with different types of property changes (4.5 months, size and shape; 7.5 months, pattern; 11.5 months, color).

These are very creative and interesting studies. However, there is another possible interpretation of the results. The narrow/wide-screen events are very similar to a phenomenon studied in mid-level vision: the tunnel effect (e.g. Burke, 1952). The basic paradigm is to show adults an object going behind a screen, or into a tunnel, followed by another object emerging out the other side. Under some circumstances (related to the speed of the trajectory, the relative sizes of the objects and the occluder, the shape of the inferred path, etc.), the object is perceived as persisting behind the occluder. This phenomenon is called 'amodal completion'. It is 'amodal' because we do not see the object behind the occluder (unlike apparent motion, which is 'modal completion', because we see the object's apparent motion). Nonetheless, in amodal completion, our visual system takes into account the various spatiotemporal parameters and yields a representation of a single object persisting through occlusion. Interestingly, as in apparent motion, the features of the objects play a minimal role in the tunnel effect. If we see a red ball going behind an occluder and a blue box emerging from the other side, we perceive it as the same object that has changed properties. 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.

W&B's interpretation of the narrow-screen studies requires that infants compare the sum of widths of two objects never seen together (22 cm) with a screen that is only 1 cm too narrow. To establish whether the narrow-screen events provide a salient violation for adults and to explore the nature of the anomaly when detected, 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). Half of the adults were shown the event with a very narrow screen of 15 cm, and half with the 21 cm narrow screen used in W&B's experiments. Virtually all of the participants in the 15 cm event spontaneously noted that something was anomalous, whereas over 60% of those in the 21 cm event described the event as if there were nothing unusual (e.g. 'A blue ball goes behind the green screen, followed by a red box coming out, and vice-versa'). More importantly, all 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 of two objects that could not fit behind the screen.

Why do adults not spontaneously notice the anomaly in the 21 cm event? We, like W&B, had adults explicitly judge whether both objects could fit behind the screen. We showed adults these events, using screen widths of 15, 18, 21, 24, and 27 cm, and asked them to rate, on a five-point scale, how sure they were that both objects could fit, without some magic trick. One part of the results fit nicely with those obtained in W&B's rating data. For screens of 15 and 18 cm, adults were certain that the objects could not fit, and for a screen of 27 cm, adults were certain that the objects could fit (close to the 25.5 cm estimate from W&B). However, for the 21 cm screen (the width of the narrow screen used in W&B), adults were not sure (mean rating 2.6), and for the 24 cm screen, they were only fairly sure (mean rating 4.1) that both could fit.

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, found 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). These data give further support to speculations in the literature that the object representations of young infants are closely related to those studied in experiments on object tracking and object files in adults (Leslie, Xu, Tremoulet & Scholl, 1998; Scholl & Leslie, 1999; Scholl & Pylyshyn, 1999; Xu, 1999). On the tunnel effect alternative, the infant did not use the property differences to infer two objects and judge that both objects could not fit behind the screen, contrary to Wilcox and Baillargeon (1998a,b) and Wilcox (1999).

Although we do not think that W&B (1998a) and Wilcox (1999)'s narrow-screen experiments show that infants use property differences to establish representations of two numerically distinct objects, we acknowledge that other studies reviewed by N&B have provided better evidence (e.g. the single trajectory event-mapping experiment, and Needham's box/cylinder experiment). Thus, the conflicting results remain. The lesson we draw from our discussion of the narrow-screen results is simply that in order to study the bases upon which infants individuate objects, we must have tasks which ask the infants more directly whether or not they have established representations of two numerically distinct objects. Event-mapping (as opposed to event-monitoring) looking time studies do so, because infants are shown outcomes of one or two objects directly on the test trials. Recent studies by Van de Walle, Carey and Prevor (1999) have developed a new manual search paradigm to address this question with a different dependent measure.

3.2. More convergent evidence – a new manual search paradigm

Consider the following task. You are shown a box into which you cannot see. You see one object (say, a toy duck) removed from and replaced in the box, and then another object (say, a toy telephone) is removed from and replaced into the box. You are then given the box – how many objects do you try to retrieve from the box? Patterns of manual search would provide direct evidence concerning object individuation. Did the infant use the differences between the duck and the telephone to form a mental model of two distinct objects in the box? What if the infant is given spatiotemporal evidence in the initial presentation – the toy duck and the toy telephone are shown simultaneously, before being returned into the box?

Besides the fact that such a task provides direct evidence concerning infants' representation of number, we had another reason for developing the manual search paradigm. N&B argued that increasing information-processing demands was responsible for the later success in our studies. One consequence of this analysis is that further increasing the information-processing demands of the tasks, relative to our earlier studies, should push the age of success to even later than 12 months.

The manual search paradigm presents infants with a much more difficult task than any experiments to date that investigate the development infants' ability to individuate objects. At no time during the test period are infants presented with any visible information to indicate how many objects are in the box. Unlike previous looking time studies, infants cannot compare a test outcome with a stored image of the objects, or a stored representation of the familiarization event. Rather, infants must create a representation of the first object, hold it in mind to compare it to a second object, then decide whether one or two objects are in the box. They must then continue to hold the representation in mind for a substantial period of time (about 20 s under some circumstances) to guide manual search. This task is substantially more difficult even than the 'event-mapping' tasks used by Xu and Carey (1996), where infants are required to map a fully visible test display onto a representation of a previously seen occlusion event. It also makes substantially more demands on short-term memory than the object segregation tasks. Further, this task requires more robust representations of the objects to support a manual search rather than a looking response, requires means-ends planning (e.g. Baillargeon, Graber, DeVos & Black, 1990) and requires prospective control of action (e.g. Bertenthal, 1996).

Van de Walle et al. (1999) contrasted a spatiotemporal condition (in which, on two object trials, infants saw both objects at the same time) with a kind/property alone condition (in which, on two object trials, they saw the two objects one at a time). The dependent measures (number of repeated reaches into the box and search time) reflected the persistence of searching for a second object (which had been surreptitiously removed) on two-object trials. Ten-month-old infants succeeded when given spatiotemporal evidence, but failed to differentiate the one- and two-object trials when given kind/property evidence alone. Twelve-month-old infants succeeded in both conditions. Thus, in an experimental paradigm that provided a more direct measure of object individuation, and made information-processing demands far greater than any previous study to date, the developmental pattern observed in Xu and Carey (1996) and Xu et al. (1999) was again observed.

Three series of studies with a wide range of information-processing demands have provided converging evidence for a shift between 10 and 12 months on the bases for successful object individuation. In addition, the series of studies described in Section 2.1 also speak of the issue whether different information-processing demands of the tasks are the major source of variance in the likelihood of success. At 12 months, infants are clearly able to process complex objects and complicated experimental procedures (as shown in the three sets of studies described above). However, in these property experiments they failed at a task with simpler objects (e.g. a red ball and a green ball). Taken together with the manual search studies, we have shown that at 12 months, infants succeeded at a task that had an increase in information-processing demands compared to Xu and Carey (1996) and failed at a task which had reduced the information-processing demands by using simpler objects. These studies provide further evidence that changes in the infant's information-processing capacity are not the only source of development.

4. Resolving the conflicts in the literature: a working hypothesis

Four series of experiments (Van de Walle et al., 1999; Xu & Carey, 1996; Xu et al., 1997, 1999) provide converging evidence for a pattern of development across a range of information-processing demands and across looking and manual search paradigms. The data from these studies suggest the presence of some developmental change that is orthogonal to the development of general memory capacity, information-processing capacity, or executive function skills. Our working hypothesis is that kind concepts (or specific sortals, for discussion see Macnamara, 1986; Xu, 1997, 1999; Xu & Carey, 1996) are acquired as infants approach their first birthday, and that the greater information-processing demands (along relevant dimensions, to be specified below) of our tasks require the infants to draw upon kind representations.

4.1. Object individuation and the distinction between property and kind representations

We begin with an analysis of object individuation in adults and the importance of distinguishing property and kind representations. Our efforts are guided by the observation that for adults, individuation occurs at multiple levels of the cognitive architecture. Three are particularly relevant to the present discussion: (1) individuation within clearly perceptual processes such as figure/ground segregation (e.g. Marr, 1982; Rock, 1975; Wallach, 1949); (2) tracking of individuals by the attentional systems that establish and maintain object files (e.g. Kahneman, Treisman & Gibbs, 1992; Scholl & Pylyshyn, 1999; Trick & Pylyshyn, 1994); and (3) individuation in terms of specific sortals such as *dog* or *car* at the conceptual level (e.g. Hirsch, 1982; Macnamara, 1986; Wiggins, 1980; Xu & Carey, 1996).

Undoubtedly, multiple sources of information are drawn upon by each of the three individuation processes identified above. The process of figure/ground segregation employs both featural information (gestalt properties such as similarity and symmetry) as well as spatiotemporal information (e.g. binocular depth cues that specify one surface as in front of another; Peterson, 1994 and references cited therein). Furthermore, recent work by Peterson and colleagues (see Peterson, 1994, for a review) shows that information about familiar shape ('experiential knowledge', see Section 4.3 below) enters into figure/ground computation at the earliest stage of processing, in parallel with featural and spatiotemporal information. The attentional mechanism of object tracking primarily uses spatiotemporal information to establish and track object files; featural/property information plays a decidedly secondary role. Leslie et al. (1998) and Scholl and Pylyshyn (1999) have argued that this mechanism underlies a body of results on object individuation and numerical competence in infancy (an insight we drew upon in our discussion of the narrow-screen studies above). Lastly, at the conceptual level, processes of object individuation and identity tracing predominately draw upon categorizing an entity as a member of a kind. Kind concepts provide criteria for individuation and identity (e.g. a cat sitting on the window-sill today and the mug sitting on the same window-sill tomorrow are two distinct objects); property/featural information, on its own, does not. Property changes are indicative of identity change only relative to kind membership, e.g. dogs can change size such that a puppy seen a month ago may be the same dog as the grown canine seen today, but cups can not. Even spatiotemporal continuity is sometimes overridden by kind considerations. For example, we judge that people cease to exist when they die, in spite of the spatiotemporal continuity of their bodies, and hundreds of encounters with two adjacent objects on our desks, say an ink bottle on top of a book, do not lead us to 'melt' the two objects into one.

This analysis requires that we distinguish between conceptual categories (kinds) and perceptual categories determined by properties (round entities, red entities, entities of a certain shape…). As Wiggins (1980) and Mandler (1999) draw the distinction, kinds provide the natural answers to 'What is it?' whereas properties do not (compare 'It is a dog' with 'It is furry'). Kind concepts are stable, long-term representations that capture information about causally correlated features of objects. Consequently, they support inductions that are not based on perceptible object properties. Yet another important aspect of kind representations for present purposes is that they are symbols for categories and individuals. Such symbols can be directly placed into short-term memory (as opposed to features within object-files) and participate in propositional representations. The format of these symbols may be mentally represented lexical items, abstract symbols in the language of thought (Fodor, 1981), or perhaps interpreted images (see, for example, Sternberg, 1975).

Armed with this conceptual analysis, we suggest that kind representations are required for object individuation under many circumstances, including when featural information is ambiguous, when spatiotemporal information is strong and needs to be overridden, and when the short-term memory demands of the task are high. As we will see, infants' successes in experiments by Baillargeon, Needham, and Wilcox, and their failures in our tasks all show that spatiotemporal information dominates other sources of information, which suggests that the object-tracking system, not kind representations, underlies object individuation for most of the first year of life.

4.2. How our working hypothesis helps resolve the conflicts in the literature

The question raised in the literature can be stated as follows. N&B reviewed studies showing that infants as young as 4 months of age sometimes use property information, including experientially derived shape representations, for object individuation. Why did 10-month-old infants (and even 12-month-olds in Xu et al., 1997) fail at object individuation on the basis of property differences?

We suggest that our tasks provide more ambiguous featural evidence than those reviewed by N&B, provide stronger spatiotemporal evidence than those reviewed by N&B, and make stronger demands on short-term memory than those reviewed by N&B. Under these circumstances, kind representations are required for adult performance in these tasks because kind representations can disambiguate featural evidence, override stronger spatiotemporal evidence, and they are symbols that can be directly placed in short-term memory.

Consider the differences between Xu et al. (1999) and Needham (1998) as an example. First, Xu et al.'s task provided stronger spatiotemporal evidence for one object. Ten-month-old infants were fully habituated to the static display of a duck perched on top of a car (or the cup on the shoe), which may have given the infant spatiotemporal evidence that there was one object. In Needham's object segregation studies, in contrast, the objects are seen static, side by side, only for a few seconds before one is grasped and moved. Second, Xu et al.'s displays were highly ambiguous regarding a purely perceptual parse, whereas the featural information in the box/cylinder displays was unambiguous. The color, texture, and shape differences in the box/cylinder displays all supported the same parse. The duck and the car (and the cup and the shoe), in contrast, were multi-parted and multi-colored. The bottom surface of the duck overlaps completely with the top of the car, the clearest discontinuity in terms of contour occurs between the head of the duck and the rest of the display, and the bill of the duck and the wheels of the car are all reasonably good candidates for objects on purely featural grounds. Therefore, this display poses a difficult problem: the spatiotemporal evidence for one object is strong and the featural information in support of the correct parse is weak. Under such circumstances, the infant must draw upon kind representations to succeed at this task.

This analysis provides an account of why complexity matters in these studies. Complexity (defined in terms of multiplicity of parts and features) makes featural information increasingly ambiguous. Also, some 'complexity' effects may not actually reflect complexity at all. Needham and Baillargeon (1998) showed that 8-month-olds succeeded in parsing the box and the straight hose into two objects when they could see the connection between the two, but failed when the hose was curved and they could not see the connection, and called this an effect of complexity. Perhaps one role of mutually supporting featural information is to focus the infant on the connecting surface of the two segments, making it more likely that they will detect the boundary. On this analysis, the important variable is the visibility of the boundary, rather than the curved/straight contrast.

Next consider the studies in Xu and Carey (1996) and compare them with the simplified event-mapping studies by Wilcox and Baillargeon (1998a). W&B showed that single trajectory (i.e. if each object was shown only once and neither object reversed its trajectory) made the task easier, such that even 7.5-month-old infants succeeded in using the difference between a box and a ball to establish representations of two objects. This ability, however, is very fragile. Multiple emergences (either one object reverses its trajectory once, or if objects were shown multiple times along a single trajectory with reversals, as in Xu & Carey, 1996) wiped out the success completely. Why should this be so? First, multiple emergences, back and forth on a single oscillating trajectory, may provide stronger spatiotemporal evidence for a single object, as in the tunnel effect or apparent motion literature. Kind representations may be needed to counteract such relatively strong spatiotemporal evidence for a single object. Second, multiple emergences place higher demands on short-term memory. Infants must keep track for each emergence, whether they have seen that object before. The availability of direct symbols for kinds, duck/car, would surely facilitate this process.

The manual search studies by Van de Walle et al. (1999) are subject to the same analysis. Ten-month-old infants' search behavior suggested that they established a representation of a single object in the box when shown two featurally distinct objects (e.g. a duck and a car) one at a time. The fact that each of the objects was taken out from the same box (considered as one location) may have provided positive spatiotemporal evidence for a single object. And the short-term memory demands of this task are the greatest of all the studies reviewed by N&B.

Lastly, the property experiments by Xu et al. (1999b) provide corroborating evidence that it is indeed kind representations that help the infant to override spatiotemporal information at 12 months. That is, the successes at 12 months cannot be explained by simply appealing to stronger featural/property representations.

As suggested by N&B, kind representations help in these cases because they are 'summary representations'. They allow the infants to encode the objects directly as a duck, a car, a cup, or a shoe, and hold such representations in short-term memory. These representations enable the infant to parse the displays according to kind membership and to override spatiotemporal evidence or ambiguous featural information.

4.3. On not identifying experiential knowledge with kind representations

N&B reported two studies showing that infants as young as 4.5 months draw upon experientially derived representations of shape in object individuation (Needham, 1999). They further suggested that these results militate against our hypothesis that kind concepts do not support object individuation until the end of the first year of life. In reply, we note that experientially derived representations of shape are not equivalent to representations of kind (Bloom & Markson, 1998; Gelman & Ebeling, 1999; Mandler, 1999; Soja, Carey & Spelke, 1991). Recent work by Peterson and her colleagues dramatically underscore this point, and is highly relevant in the present context. These studies show that experientially derived shape representations contribute to object individuation at very early stages of perceptual processing and that these representations are dissociable from those implicated in recognition of kind membership.

Peterson's work concerns the processes that establish figure/ground representations. Establishing a representation of figure is a paradigmatic individuation problem, for a single figure is seen as an individual with a defined shape in front of an undifferentiated background. As is well known, both featural information (gestalt properties of arrays, such as similarity, symmetry, boundedness) and spatio-temporal information (e.g. binocular depth cues that specify one surface as in front of another), enter into the computations that resolve figure/ground ambiguity at the earliest stages of processing. What Peterson has shown is that 'experiential' knowledge also enters into the computations, and does so at the earliest stages of processing. In a series of studies, she has compared figure/ground displays in which one of the surfaces is bounded by a meaningful shape (e.g. a face profile or a sea horse; see Fig. 1), 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 other factors at the very earliest stages of figure determination (e.g. within the first 14 ms). That is, the meaningful shape is more often seen as figure than its complement, and this factor sometimes overrides other cues to figure such as symmetry or depth cues (e.g. Peterson & Gibson, 1993).

Most relevant to us here, Peterson et al. (1999) presented neuropsychological evidence that the experientially derived shape representations that enter into this process are not the kind representations that mediate object recognition. They presented a double dissociation between a visual agnosic patient with bilateral

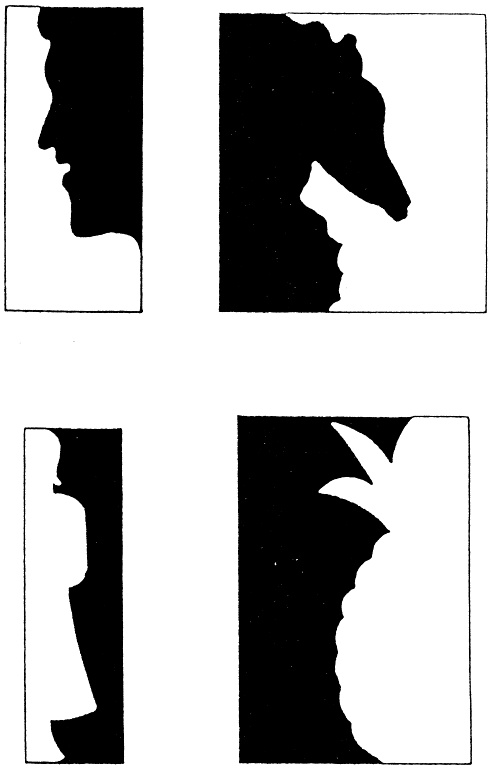

Fig. 1. The figure/ground stimuli used in Peterson and Gibson (1993).

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.

In sum, Peterson's research demonstrates that kind representations (those involved in placing a stimulus in a meaningful conceptual category) are not the only experientially derived shape representations that play a role in processes of individuation. Thus, Needham's elegant experiments demonstrating that very young infants bring experiential information to bear on object individuation do not necessarily show that these infants represent kind concepts. This is clearly an open question that awaits further research. Other relevant research, however, suggests that it is unlikely that the young infants in Needham's studies drew on kind representations. We will conclude our comments with a brief review of other relevant research consistent with our hypothesis that kind concepts are not acquired until the last few months of the first year.

5. Reasons to believe that kind representations emerge towards the end of the first year

Two other sources of data are consistent with the claim that kind representations emerge at the end of the first year of life: experiments concerning the role of labeling in object individuation and categorization, and experiments concerning the basis of inductive inference in infancy.

5.1. The role of labeling in categorization and object individuation

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 property or kind (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 in 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.

Studies by Waxman and her colleagues (Balaban & Waxman, 1996; Waxman & Markow, 1996) provided evidence that labeling facilitates categorization in 9- and 13-month-old infants. Furthermore, Waxman (1999) showed that by 13 months, infants distinguish between kind representations and property representations. Upon hearing a series of objects described by a count noun (‘Look, it is a blicket’), they extract kind similarity (at both the basic and superordinate levels), whereas upon hearing an adjective (‘Look, it is a blickish one’), they also extract property similarity (texture and color).

5.2. Inductive inference in infancy

In a series of elegant studies, Mandler and her colleagues have provided evidence that by 9–11 months of age, infants can form categories of objects that are not exclusively driven by perceptual similarity, but rather seem to be conceptual in nature (for a review see Mandler, 1999). For example, 11-month-old infants’ categorization of animals, vehicles, kitchen utensils, plants, and furniture seems to be driven by global kind category membership rather than overall perceptual similarity (Mandler & McDonough, 1993, 1996). Indeed, infants this age even appear to be capable of making simple inductions on the basis of membership in the category animal or vehicle, one of the hallmarks of conceptual representations of object kinds (Mandler & McDonough, 1996). Infants generalize actions, such as keying a car to other vehicles, and putting a dog to bed to other animals, irrespective of perceptual similarity (see Mandler, 1999). Similarly, Baldwin, Markman and Melartin (1993) have demonstrated that the capacity to project a non-obvious functional property from one object to another of the same similarity emerges robustly between the ages of 9 and 12 months.

6. Conclusion

In this rejoinder, we have laid out some areas of agreement and disagreement with Needham and Baillargeon and sketched why we are not convinced by their arguments. A full resolution of the conflicting data in the literature requires further conceptual and empirical work. Still, our working hypothesis that genuine kind representations emerge towards the end of the first year, and that younger infants fail just in case kind representations are required for adult performance on some tasks, does a reasonable job in accounting for all the data to date. We admire Needham, Wilcox, and Baillargeon's extremely elegant research programs and applaud their efforts to bring this debate to the attention of the cognitive science community at large. This has been a very fruitful exchange, and we look forward to more of them with Needham, Wilcox, and Baillargeon as well as with other researchers.

As we have argued above, object individuation occurs at different levels of processing. It is an extremely challenging task for cognitive scientists to draw the right architectural distinctions within adult information processing, making the task of charting the developmental course of distinct mechanisms all the more difficult. Although it is crucial to discover the various factors which influence infants' performance on different tasks, we need to go one step further. We need to be guided by studies on adults' cognitive architecture in order to locate the phenomena we have discovered, so we may specify where development occurs and how.

Acknowledgements

This research was supported by an NIH B/START grant (1R03 MH59040-01) to Fei Xu and an NSF grant (SBR-9514695) to Susan Carey. We thank Cristina Sorrentino, Elizabeth Spelke, and Gretchen Van de Walle for many helpful discussions and/or comments on an earlier draft of the paper. Mistakes remaining are, of course, entirely ours.

References

Baillargeon, R., Graber, M., DeVos, J., & Black, J. (1990). Why do young infants fail to search for hidden objects? *Cognition*, *36*, 255–284.

Balaban, M., & Waxman, S. (1996). Words may facilitate categorization in 9-month-old infants. *Journal of Experimental Child Psychology*, *64*, 3–26.

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infants' ability to draw inferences about nonobvious object properties: evidence from exploratory play. *Child Development*, *64*, 711–728.

Bertenthal, B. I. (1996). Origins and early development of perception, action, and representation. *Annual Review of Psychology*, *47*, 431–459.

Bloom, P., & Markson, L. (1998). Intention and analogy in children's naming of pictorial representations. *Psychological Science*, *9*, 200–204.

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

Carey, S. & Bassin, S. (1998). *The violation of expectancy method: when adults fail to see the trick*. Poster presented at the International Conference for Infant Studies, Atlanta, GA.

Fodor, J. A. (1981). *The language of thought*, Cambridge, MA: Harvard University Press.

Gelman, S. A., & Ebeling, K. S. (1999). Shape and representational status in children's early naming. *Cognition*, *66*, B35–B47.

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

Johnson, S., & Aslin, R. (1996). Perception of object unity in young infants: the roles of motion, depth, and orientation. *Cognitive Development*, *11*, 161–180.

Jusczyk, P. W., Johnson, S. P., Spelke, E. S., & Kennedy, L. J. (1999). Synchronous change and perception of object unity: evidence from adults and infants. *Cognition*, *71*, 257–288.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology*, *24*, 174–219.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*, 483–524.

Kellman, P. J., Spelke, E. S., & Short, K. (1986). Infant perception of object unity from translatory motion in depth and vertical translation. *Child Development*, *57*, 72–86.

Kellman, P. J., Gleitman, H., & Spelke, E. S. (1987). Object and observer motion in the perception of objects by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 586–593.

Leslie, A., Xu, F., Tremoulet, P., & Scholl, B. J. (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. (1999). Perceptual and conceptual processes in infancy. *Journal of Cognitive Development*, in press.

Mandler, J., & McDonough, L. (1993). Concept formation in infancy. *Cognitive Development*, *8*, 291–318.

Mandler, J. M., & McDonough, L. (1996). Drinking and driving don't mix: inductive generalization in infancy. *Cognition*, *59*, 307–335.

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

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

Needham, A. (1999). Infants' perception of a paintbrush and a keyring: use of categorical versus configural knowledge to segregate a display, manuscript in preparation.

Needham, A., & Baillargeon, R. (1998). Object segregation in 8-month-old infants. *Cognition*, *62*, 121–149.

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

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

Peterson, M. A., de Gelder, B., Rapcsak, S. Z., Gerhardstein, P., & Bachoud-Levis, A. (1999). A double dissociation between conscious and unconscious object recognition processes revealed by figure-ground segregation. *Vision Research*, in press.

Rock, I. (1975). *An introduction to perception*, New York: MacMillan.

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

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

Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, *10*, 253–270.

Soja, N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition*, *38*, 179–211.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*, 29–56.

Spelke, E. S., Breinlinger, K., Jacobson, K., & Phillips, A. (1993). Gestalt relations and object perception: a developmental study. *Perception*, *22*, 1483–1501.

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

Sternberg, S. (1975). Memory scanning: new findings and current controversies. *Quarterly Journal of Experimental Psychology*, *27*, 1–32.

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

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

Van de Walle, G., Prevor, M. & Carey, S. (1999). The use of kind distinctions for object individuation: evidence from reaching, submitted.

von Hofsten, C., & Spelke, E. S. (1985). Object perception and object-directed reaching in infancy. *Journal of Experimental Psychology: General*, *114*, 198–212.

Wallach, H. (1949). Some considerations concerning the relationship between perception and cognition. *Journal of Personality*, *18*, 6–13.

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

Waxman, S., & Markow, D. (1996). 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*, in press.

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

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

Xu, F. (1997). From Lot's wife to a pillar of salt: evidence that *physical object* is a sortal concept. *Mind and Language*, *12* (3/4), 365–392.

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

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

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

Xu, F., Carey, S. & Quint, N. (1997). *Object individuation at 12 months: shape and other properties*. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC, manuscript in preparation.